Electron Tomography Image Reconstruction Using Data-Driven Adaptive Compressed Sensing

Ala' Al-Afeef,[1,2] W. Paul Cockshott,[1] Ian MacLaren,[2] and Stephen McVitie[2]

[1]School of Computing Science, University of Glasgow, Glasgow, United Kingdom
[2]SUPA School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom

Summary: Electron tomography (ET) is an increasingly important technique for the study of the three-dimensional morphologies of nanostructures. ET involves the acquisition of a set of two-dimensional projection images, followed by the reconstruction into a volumetric image by solving an inverse problem. However, due to limitations in the acquisition process, this inverse problem is ill-posed (i.e., a unique solution may not exist). Furthermore, reconstruction usually suffers from missing wedge artifacts (e.g., star, fan, blurring, and elongation artifacts). Recently, compressed sensing (CS) has been applied to ET and showed promising results for reducing missing wedge artifacts. This uses image sparsity as *a priori* knowledge to improve the accuracy of reconstruction, and can require fewer projections than other reconstruction techniques. The performance of CS relies heavily on the degree of sparsity in the selected transform domain and this depends essentially on the choice of sparsifying transform. We propose a new image reconstruction algorithm for ET that learns the sparsifying transform adaptively using a dictionary-based approach. We demonstrate quantitatively using simulations from complex phantoms that this new approach reconstructs the morphology with higher fidelity than either analytically based CS reconstruction algorithms or traditional weighted back projection from the same dataset. SCANNING 38:251–276, 2016. © 2015 Wiley Periodicals, Inc.

Key words: electron tomography, image reconstruction, PTB_7:$PC_{71}BM$ blends, compressed sensing, solar cells

Contract grant sponsor: Lord Kelvin Adam Smith Scholarship.

Address for reprints: Ala' Al-Afeef, School of Computing Science, University of Glasgow, Glasgow G12 8QQ, United Kingdom.
E-mail: a.al-afeef.1@research.gla.ac.uk

Introduction

Transmission electron microscopy (TEM) is a key technique that is widely used for characterising nanostructures. However, when a morphological analysis of complex materials is required, the fact that the images are 2D projections of a 3D object can make interpretation difficult. Whilst stereological techniques have been used for more than 40 years, a more quantitative approach using electron tomography (ET) is now widely used and provides a better solution to visualising and understanding 3-dimensional nanostructures (De Rosier and Klug, '68; Baumeister *et al*., '99). ET uses TEM to collect a sequence of possible 2D projection images at different tilts, followed by alignment and reconstruction to obtain a 3D volume. Until recently, the 3D reconstruction was typically carried out using the weighted back-projection (WBP) algorithm or an iterative reconstruction algorithm such as the simultaneous iterative reconstruction technique (SIRT) (Gilbert, '72). Nevertheless, such reconstruction algorithms have been known for some time to generate serious artifacts, and there has been significant recent work on alternative reconstruction methods.

In recent years, ET has been widely used in nanoscience and has yielded important results inaccessible by conventional 2D projection imaging alone (for instance Arslan *et al*., 2005; Lucic *et al*., 2005; Al-Amoudi *et al*., 2007; Robinson *et al*., 2007; Arslan and Stach, 2012; Scott *et al*., 2012). In practical ET, the quality of 3D reconstructed tomograms depends on many different factors, principally the maximum angular range and the number of acquired projections. The angular range is usually limited due to unavoidable technical restrictions such as the limited space between the objective lens pole-pieces of the electron microscope or the increasing thickness of the specimen at high-angle tilts which makes it no longer electron transparent (Williams and Carter, 2009), although this can be circumvented in some cases where the sample can be prepared as a needle specimen (e.g., Midgley and Dunin-Borkowski, 2009). The number of projections possible for a sample can be limited, especially for

beam-sensitive materials, in order to keep the total electron dose below some critical level. Consequently, two critical problems arise (Midgley and Weyland, 2003; Kawase *et al*., 2007; Midgley and Dunin-Borkowski, 2009):

1. Missing wedge artifacts: elongation and blurring of the reconstructed object in the direction of the optical axis due to a missing wedge of information caused by limited angular range.
2. Resolution degradation: the lower the number of projections, the lower the reconstruction resolution (Radermacher, '88).

The quality of a tomographic reconstruction can be enhanced by including additional prior knowledge about the specimen throughout the reconstruction process. A relatively recent prior knowledge technique, compressive sensing (CS) (Candès *et al*., 2006; Donoho, 2006; Candès and Wakin, 2008) has been applied with great success in magnetic resonance imaging (MRI) (Candès *et al*., 2006; Lustig *et al*., 2007a). CS has more recently been applied to ET (Saghi *et al*., 2011; Binev *et al*., 2012; Goris *et al*., 2012; Monsegue *et al*., 2012; Leary *et al*., 2013) and image acquisition for high-resolution scanning TEM (Stevens *et al*., 2013). It has been demonstrated that even with reduced dataset, it is possible to reconstruct tomograms with high fidelity and reduced missing wedge artifacts (Saghi *et al*., 2011). Such advantages make CS an effective method for decreasing beam damage, obtaining reliable, high-resolution morphology, and enabling quantitative measurements from 3D tomograms. The key prior knowledge employed in CS is that the signal is likely to be sparse in a transform domain. If such a sparsifying transform can be determined, the original signal can then be accurately reconstructed from a set of measurements significantly sparser than that which would be required by the Nyquist sampling theorem (Candès *et al*., 2006; Donoho, 2006; Candès and Wakin, 2008; Fornasier and Rauhut, 2011). In terms of Shannon information theory, compressive sensing exploits signal redundancy.

As sparsity is the key requirement, to derive any benefit from compressive sensing; researchers have experimented with a variety of sparsifying transforms (Tsaig and Donoho, 2006; Gan, 2007; Lustig *et al*., 2007a]. The transforms hold for signals that are sparse in the standard coordinate basis or other orthonormal basis. One common sparsifying transform is the total variation (TV), which is a summation of the discrete gradient transform (DGT) coefficients. In the case of ET, CS algorithms have been suggested for tomographic reconstruction by maximising sparsity in the TV domain. For example, Goris *et al*. (2012) have investigated the use of a TV minimisation (TVM) in ET. Saghi *et al*. (2011) proposed a CS-ET algorithm, which maximises sparsity in both the TV and spatial domains. Also, they showed that CS-ET was able to reconstruct 3D maps from a very limited number of tilt images. Furthermore, Monsegue *et al*. (2012) showed that elongation artifacts caused by limited angle sampling can be effectively reduced using anisotropic TV (i.e., decreasing contributions from the missing wedge direction).

In spite of the success of these algorithms, they are applicable to signals that are sparse in certain predefined (fixed) sparsifying transforms. In many practical examples, the signal under study is not compressible (sparse) in such transforms. For example, TV minimisation can be effective for reconstruction if the object under study can be described as a piecewise constant (i.e., having sharp boundaries); however, this is not true for many samples. Other drawbacks of using the TV operator include over-smoothing of fine structures, difficulties in separating true structures from noise, and a degradation of spatial resolution (which becomes especially apparent in noisy examples). Consequently, there are compelling reasons to investigate alternative sparsifying methods for CS-based ET reconstruction in order to avoid such difficulties.

The choice of a sparsifying transform is typically decided using some simplifying assumptions, usually utilising a pre-chosen basis set such as steerable wavelets or curvelets. An alternative effective approach is by learning the sparsifying transform directly from examples (i.e., images) adaptively. The ground-breaking work of Olshausen ('96) was the first to suggest an algorithm in the field of natural image processing for learning a transform to find sparse linear codes for natural scenes under the sparse approximation assumption. The learned transform can thus be adapted for specific types of training images. This enhances the range of images that can be compressed. Sparse representation with learned transforms outperforms predefined transforms in a range of image processing applications like de-noising, de-blurring, and in-painting (Aharon *et al*., 2006b; Elad and Aharon, 2006; Mairal *et al*., 2008; Protter and Elad, 2009). Furthermore, the recent work by Liu *et al*. (2014) showed the feasibility of a dictionary-based statistical reconstruction approach for discrete tomography at atomic resolution, using a simulated binary phantom that mimics rows of discrete atoms.

A recent proposal (Rauhut *et al*., 2008; Candès *et al*., 2011) extended CS theory to the common situation where signals are not sparse in an orthonormal basis but rather in overcomplete sparsifying transforms (also known as redundant dictionaries). The use of overcomplete dictionaries gives more flexibility to represent images, especially when there are no known good sparsifying transforms. Sparsity can be enhanced by redundancy as dictionaries will contain more effective atoms to capture various features of an image. Rauhut *et al*. (2008) discovered that if a given signal is sparse using an overcomplete/redundant dictionary, and

the restricted isometry property (RIP) (Candès *et al.*, 2006) is satisfied for a combination of that dictionary and a sampling matrix, then it is possible to implement CS by an existing CS optimisation algorithm, such as Basis Pursuit (Chen *et al.*, '98). This work opens up the possibility of handling the tomography problem as a direct sparse decomposition method using redundant dictionaries (Starck *et al.*, 2004), which can significantly help in reducing artifacts.

By combining the concept of redundancy and sparsity with dictionary learning, the sparsity of a signal can be further enhanced, and the recovery of a good approximation of the original signal can be achieved from fewer measurements with higher fidelity. Furthermore, when dictionary training is performed using small overlapping regions (patches) that are extracted from the image under reconstruction, an additional averaging effect arises. This averaging effect can be efficient in denoising and suppressing down-sampling artifacts without losing resolution (Elad and Aharon, 2006).

In this article, we extend our work in (AlAfeef *et al.*, 2014,2015a,b) and explore the strengths of patch-based adaptive redundant dictionaries for electron tomography. The main contribution of this article is the development of a fast and robust numerical algorithm, which we have named DLET, for ET reconstruction and the application on experimental dataset. The algorithm employs a sparse reconstruction technique that incorporates prior information by means of adaptive dictionaries. The dictionary is adapted to the data and is learned during the reconstruction process in a way that leads to a sparser representation of the underlying image. The technique is tested using a simulated phantom in both noisy and noiseless cases and compared to other techniques utilising fixed sparsifying transforms in ET. We also provide a comparative study using experimental tilt series of polymer solar cells.

This article is organised as follows: Background Section discusses dictionary learning and compressed sensing for ET reconstruction. Methodology Section presents the method and algorithm for ET reconstruction from the under-sampled data. Following this, Numerical Simulation Section presents the numerical experimental results and analysis for two case studies followed by an experimental case study using projection data obtained using bright field transmission electron microscope imaging of PTB7:PC71BM polymer blends (for solar cell applications) in Three-Dimensional Morphology of Organic Photovoltaic (OPV) Solar Cells Section. Finally, we conclude with a brief discussion.

Background

The reconstruction process of ET involves generating 3D volume data from several 2D projections of a real space *object f(x, y, z),* viewed from different angles. Since the 2D projections in ET are acquired in single axis parallel beam geometry (i.e., tilting around the y-axis while the electron beam is parallel to the microscope optical axis [as in Fig. 1]), the volume reconstruction problem can be simplified by reconstructing each 2D slice $f(x, y = SectionNumber, z)$ from the corresponding 1D projections. The process of projecting $f(x, const, z)$ with increment ds along lines L determined by a tilt angle β and the distance γ to the origin, is referred to as the Radon transformation R_{yi} (Radon, '36):

$$R_{y_i}(\gamma,\beta) = \int_L f(x,y_i,z)ds \qquad (1)$$

The discrete outcome of this projecting process is called a *sinogram*. Thus, the reconstruction of a sinogram can be

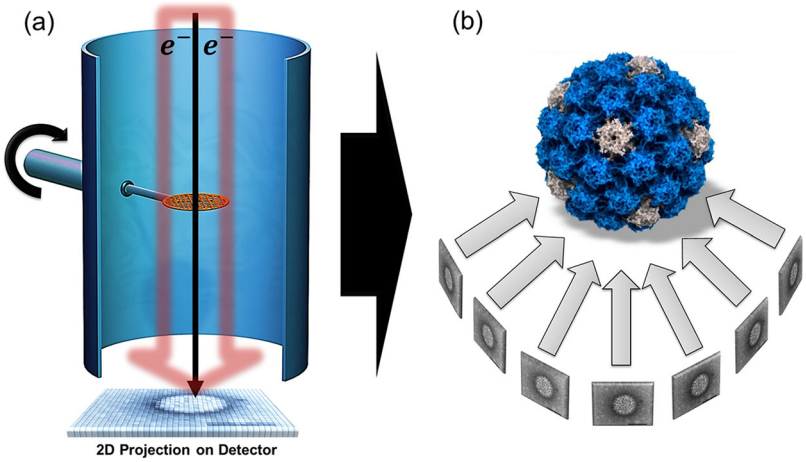

Fig 1. Schematic illustrating the data collection and reconstruction principles for electron tomography: (a) The specimen is tilted relative to the electron beam and a tilt series of projection images is acquired. (b) The tilt series are then aligned and reconstructed to generate a three-dimensional tomogram of the specimen. This tomogram can be segmented to generate a 3D model for quantitative study.

mathematically described as the inversion of the Radon transform (R^{-1}). In principle, given a sufficient number of projections, the tomographic reconstruction of an object can be obtained by applying the weighted back-projection method. Of course, this requires that the imaging produces a signal that is truly proportional to the thickness of the object, and if the imaging process is non-linear (e.g., intensity is a decaying exponential function of thickness), then some pre-processing of the images may be required before inversion of the Radon transform is sensible or possible (van den Broek *et al*., 2012).

Tomographic reconstruction can also be implemented in the Fourier domain using the direct Fourier inversion technique (Kak and Slaney, '88). This approach is based on the central slice theorem (Cramer and Wold, '36; Ramachandran and Lakshminarayanan, '71; Natterer and Ẅubbeling, 2001) stating that the Fourier transform of an object's projection at a given angle is a central plane through the Fourier transform of the object (Deans, '83). Hence, by taking 1D Fourier transforms of projections at many different angles, many Fourier slices will be sampled. In theory, tomographic reconstruction is achievable from the inverse 2D Fourier transform. However, the inversion in either real space or Fourier methods is not straightforward since the inverse reconstruction problem is under-determined. This happens due to the missing wedge and the limited number of projections. Additionally, the reconstruction will be further degraded due to the presence of noise and alignment errors.

The quality of reconstruction can be further enhanced by incorporating further knowledge about the specimen, such as filtering regularisation and projection errors, as in SIRT (Kak and Slaney, '88; Herman, 2009), gray levels as in DART (Batenburg *et al*., 2009), prior knowledge about the shape (geometric prior knowledge) as in object-based reconstruction methods (Alpers *et al*., 2013), or sparsity of the signal as in CS-based reconstruction methods (Saghi *et al*., 2011; Leary *et al*., 2013). Here, we demonstrate the use of prior knowledge about signal sparsity to overcome the problems arising from under-sampling and the missing wedge.

Compressive Sensing Theory

Compressive sensing (CS), introduced by Donoho and Tanner (2005) and Candès *et al*. (2006), is one of the major advances in signal processing in the last decade (Candès and Wakin, 2008; Duarte and Eldar, 2011; Foucart and Rauhut, 2013). Subject to appropriate conditions, it allows one to recover useful information from far fewer measurements than is classically considered possible (namely, Nyquist sampling theorem). CS had been applied successfully in a variety of fundamental applications such as medical imaging (Lustig *et al*., 2007b, 2008), compressed imaging (Duarte *et al*., 2008), radar (Baraniuk and Steeghs, 2007), communications (Cotter and Rao, 2002), robotics (Mostofi, 2011), astronomy (Bobin *et al*., 2008) and quantum information processing (Gross *et al*., 2010). CS employs principles of sparse approximation that is well established in image and audio compression standards such as JPEG and MP3. In the context of ET, the reconstruction problem can be modelled, in the presence of noise, as:

$$B = R_\theta x + e \tag{2}$$

where $x \in \mathrm{IR}^N$ is a real-valued vector of size N in the real space IR and represents the unknown tomogram (2D image representing a slice through 3D object). $R_\theta \in \mathrm{IR}^{M\times N}$ is the discrete Radon transform operator that transforms x into projection measurements $B \in \mathrm{IR}^M$ in the sinogram domain and e is an error term modelling measurement errors. B is undersampled when $M < N$ (i.e., $R_{\theta<180^\circ}$) by one or several orders of magnitude indicating severe undersampling. Sampling is non-adaptive as R does not depend on x.

Suppose the unknown tomogram x is sparse (or approximately s-sparse) in a domain denoted by Ψ which is a linear operator that transforms from pixel representation into a sparse representation. The representation of x is said to be s-sparse in the Ψ domain if x can be represented by at most s non-zero coefficients and $s \ll N$. If Ψ captures only the most important information about x by $s \ll N$ non-zero coefficients, then x is said to be compressible in Ψ. Let R_u^f be the sampling matrix (i.e., under-sampled Radon transform implemented using the direct Fourier inversion technique discussed earlier). The standard theory of Compressed Sensing asserts that x may be recoverable from the under-sampled measurements B given that: x can be compressed by Ψ, and both R_u^f, Ψ are incoherent (dissimilar) (Candès *et al*., 2006; Donoho, 2006). The incoherence requirement is to ensure that each measurement in B provides information about many coefficients of x and to ensures that the encoding of the x coefficients is different for each measurement in B. In the tomography problem, incoherence has been proven empirically as in Lustig *et al*. (2007b). They show that in incoherent sampling, the undersampling (aliasing) artefacts should spread throughout the sparse domain, such that they appear as noise interference. Such incoherent aliasing will enable the recovery of the significant coefficients that stick above the interferences by a non-linear optimisation process.

It should be noted that despite the lack of mathematical justification for the principle of incoherence in many real-world inverse problems including tomography, CS has been, and continues to be used with great success in many areas (e.g., Lustig *et al*., 2007b, 2008; Larson *et al*., 2011). Furthermore, recent studies suggest that the inverse problem need not be incoherent and sparse, but asymptotically incoherent and asymptotically sparse

(Adcock *et al.*, 2013). This more relaxed condition can be satisfied in many inverse problem applications, including ET, and can narrow the gap between the theoretical and applied sides of the field.

The typical CS model formulation for ET can be expressed as a constrained optimization problem:

$$\min\|\Psi x\|_1 \; s.t. \; \|R_u^f x - B\|_2 \leq \varepsilon \qquad (3)$$

where B is the set of under-sampled measurements, ε is the discrepancy level that controls the fidelity of the reconstruction to the measured data and is usually set above the expected noise level, and Ψ is the sparsifying transform which is usually chosen as an orthonormal transform for images. For instance, Ψ can be a wavelet transform, which transforms x into wavelet coefficients Ψx, that are mostly zero (sparse). This objective function promotes sparsity (Chen *et al.*, '98; Fornasier and Rauhut, 2011) as it minimises the l_1 norm (sum of absolute values) of Ψx, and enforces data consistency using the constraint $\|R_u^f x - B\|_2 \leq \varepsilon$. In other words, among all solutions that are consistent with the measured data, equation (3) tries to find a simple sparse solution that is compressible by the transform Ψ.

A number of methods can be used to solve this minimisation problem. These include linear programming methods such as basis pursuit (BP) (Chen *et al.*, '98; Donoho *et al.*, 2006) (which is effective if measurements are noisy), the least absolute shrinkage and selection operator (LASSO) method (Tibshirani, '96), the focal under-determined system solver (FOCUSS) (Gorodnitsky and Rao, '97), sparse Bayesian learning (Wipf and Rao, 2004), and other sparse approximation algorithms (Rice, 2014). In many real-world applications, these algorithms usually provide better performance than predicted by existing theories based on analytical mathematics (Adcock *et al.*, 2013). Alternatively, an approximate solution can be obtained using greedy algorithms, such as orthogonal matching pursuit (OMP) (Tropp and Gilbert, 2007), which are based on replacing the l_1 norm with l_0 quasi-norm (the count of non-zero elements). A recent review of the sparse coding algorithms can be found in Tropp and Wright (2010). For more details on CS theory, we direct the reader to a more extensive coverage in Davenport *et al.* (2012) and the CS resources: http://dsp.rice.edu/cs.

The success of compressed sensing depends critically on sparsity (Candès and Wakin, 2008). In equation (3), the compression degree of the sparsifying transform Ψ is crucial to achieving high-quality reconstructions. Ψ can be either a fixed or adaptive transform. Fixed transforms have been used successfully in many applications; however, their compression degree may be limited, and in many cases, cannot be applied to any image without satisfying certain pre-conditions. For example, finite-differences (commonly known as TV) performs well for piecewise-constant images since the uniform regions and discrete boundaries can be well recovered. However, for non-piecewise-constant images, the cartoon effect can be damaging for the reconstruction quality. Therefore, other sparsifying techniques need to be considered for such images.

Furthermore, fixed transforms can produce undesirable artifacts such as blocking, blur, ringing, or edge artifacts. For example, the use of TV in image restoration can damage fine details, and cause staircase artifacts (Chan *et al.*, 2005). The use of the discrete cosines transform (DCT) or wavelet-based coding can produce ringing artifacts in an image which appears like ripples or oscillations around sharp contours or edges in the spatial domain (Marziliano *et al.*, 2004). This is caused by truncation of the high-frequency transform coefficients in DCT or wavelet-based coding.

Adaptive sparsifying techniques, on the other hand, benefit from the intrinsic sparsity that images usually have, which improves recovery by tailoring the best sparsifying basis (i.e., by dictionary learning). The learning process can be controlled in a way that produces dictionaries with improved sparsifying ability. Recent studies in image processing showed the feasibility of adaptive transform techniques for a variety of problems (Elad and Aharon, 2006; Mairal *et al.*, 2008; Protter and Elad, 2009), with state of the art results and possible extensions to inverse problems (Peyré, 2011).

Dictionary Learning

Dictionary learning (DL) is the process by which a dictionary adapted to the data is produced. Given a vector $Y \in \mathrm{IR}^N$ to represent a training image of size $N_x \times N_y = N$ pixels as in Figure 2, $M_i \in \mathrm{IR}^{n,S}$ is an operator that extracts Y into S blocks (patches) of size $\sqrt{n} \times \sqrt{n}$ pixels represented as $M_i Y = \{y_{i,j}\}_{i=1,j=1}^{S,n}$. These patches are extracted with overlaps and 1-pixel sliding, meaning that the value of each pixel in Y will be included in a maximum of n patches. The maximum number of patches in a training set $S = (N_x - \sqrt{n} + 1)(N_y - \sqrt{n} + 1)$. A dictionary is a matrix $D \in \mathrm{IR}^{n,K}$ (Fig. 3(a)) with columns $d_{i=1,2,\ldots,K} \in \mathrm{R}^{n \times 1}$ called Atoms; these form the basis that is used to approximately represent a given image. D can be initiated by selecting K patches from $M_i Y$. The choice of Y can either be the current tomogram under reconstruction or a high-quality tomogram of a similar specimen. Let $D \in \mathrm{IR}^{n,K}$ be an overcomplete dictionary by which we mean there are more entries in the dictionary than the required number of linearly independent vectors, the sum of which could be used to exactly represent any given vector as demonstrated in (Fig. 3(b)). As this always can be done with exactly n linearly independent vectors, this means that the number of entries in the dictionary $K > n$. Suppose an image patch y_i can be represented exactly or

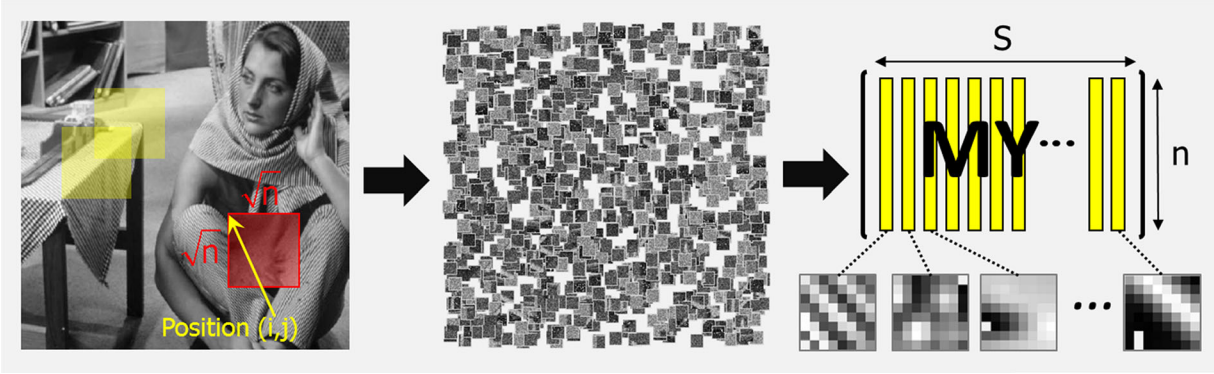

Fig 2. A schematic illustration of patch-based processing of images. By breaking the given image Y into overlapping (small) patches to be used in dictionary learning and sparse coding.

approximately as a sparse linear combination of the atoms (as in Fig. 4) drawn from drawn from D (i.e., $y_i \approx D\alpha_i$) where $\alpha_i \in \mathrm{IR}^K$ is a vector with very few nonzero entries ($\ll n$), typically around five atoms are used. Then, the DL process aims to find a possible optimal dictionary for sparse representation of training samples $M_i Y$. It can be expressed as:

$$\min_{D,\alpha} \sum_i \|M_i Y - D\alpha_i\|_2^2 \; s.t. \; \forall i \|\alpha_i\|_0 \leq n_{nz} \qquad (4)$$

where n_{nz} is the maximum number of atoms that can be used in sparse representation (i.e., the number of non-zero coefficients). The objective function in equation (4) is called orthogonal matching pursuit (OMP) for approximating the fit of a linear model with constraints imposed on the number of non-zero coefficients (i.e., the L_0 quasi-norm; Figs. 3(b) and (5)) is an example of learned dictionary. In terms of computational complexity, a problem using the L_0 quasi-norm (as in equation 4) is shown to be NP-hard (Natarajan, '95). Different techniques have been developed to solve the DL problem in (4). These can be classified into three main categories (Tosic and Frossard, 2011): (i) probabilistic learning methods such as maximum likelihood methods (Olshausen and Field, '97) and the optimal directions (MOD) method (Engan *et al.*, '99); (ii) clustering or vector quantisation methods, such as Lloyd's algorithm (Lloyd, '82) and its generalisation, the K-SVD method (Aharon *et al.*, 2006a); and (iii) learning dictionaries with a particular construction, typically driven by priors on the structure of the data (Sallee and Olshausen, 2002; Mailhé *et al.*, 2008). A review of dictionary learning methods can be found in Tosic and Frossard (2011).

Among these methods, the vector quantisation (VQ) approach has been applied successfully in many fields such as in image compression and 3D-microscopic image coding (Cockshott *et al.*, 2003; Tao and Cockshott, 2004). In particular, the K-SVD algorithms have attracted increasing interest recently because of their numerical efficiency and fast convergence (Elad, 2010; Peyré, 2011). The K-SVD algorithm typically iterates between two steps:

1. The Sparse Coding step: the minimisation problem in (4) is solved with respect to α_i with at most n_{nz} non-zero coefficient and fixed D. This step can be performed using any pursuit algorithm; however, the OMP method (Tropp and Gilbert, 2007) is typically used.

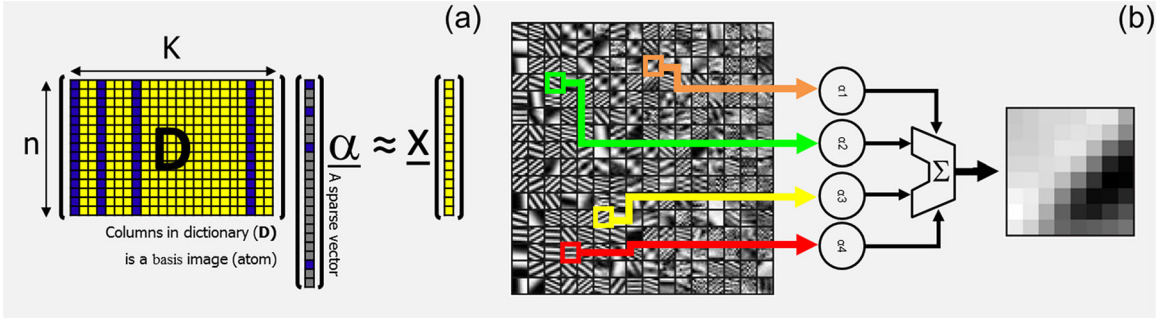

Fig 3. An example illustrating the sparse coding process, which is the problem of taking a large input image, and finding an approximate decomposition of that image using a linear blend of small images (patches) of commonly occurring subpatterns from a dictionary. (a) Sparse coding with $D\alpha_i \approx x_i$, (b) the corresponding graphical illustration. The vector α_i is sparse with $K = 4$ nonzero coefficients; the image patch (vector x) is a linear combination of these columns (images patches).

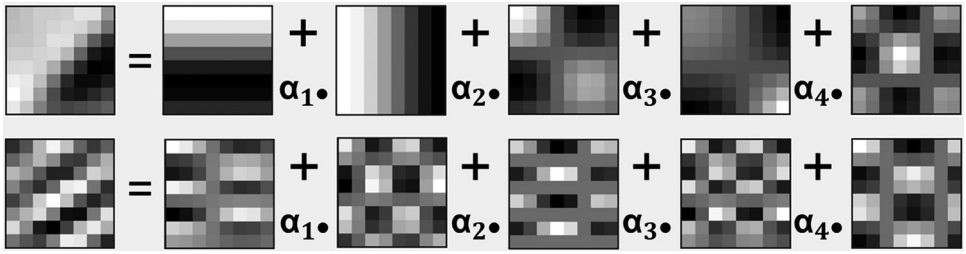

Fig 4. The illustration of Sparse coding from basis to approximate two patterns. The coefficients α_i are all of the same order of magnitude and are not shown due to space limitations.

2. The Dictionary Codebook Update Stage: the columns (atoms) of D are updated sequentially, as well as the corresponding relevant coefficients α_i.

This algorithm continues to iterate between the two steps until it converges, providing a normalised dictionary D (i.e., each atom has a unit norm). Further details about the K-SVD algorithm can be found in Aharon *et al*. (2006a). An in-depth review of the applications of dictionary learning to image processing and computer vision can be found in (Elad *et al*., 2010; Wright *et al*., 2010).

Methodology

In this work, we propose a dictionary learning compressive sensing-based technique for the ET image reconstruction problem. We call our technique dictionary learning for electron tomography (DLET). Our technique relies on the prior innovations of Lustig *et al*. (2007b) for CS-MRI, and Elad and Ahoron (2006) for dictionary learning. This article shows that 3D reconstruction quality in ET can be further enhanced when enforcing sparsity using an adaptive dictionary. To solve the ET tomographic reconstruction problem, we have used an over-complete dictionary as a regularisation term using the following formulation:

$$x' \triangleq \min_{x,D,\alpha} \frac{1}{S}\sum_{i}^{s}\frac{1}{2}\|M_i x - D\alpha_i\|_2^2 - +\lambda\|\alpha_i\|_1 s.t. \|Rx - B\|_2^2 \leq \varepsilon \tag{5}$$

where $\lambda > 0$ is the sparsity Lagrange multiplier, ε controls the discrepancy level to the measured data. This cost function enforces sparsity in the image domain x using an adaptive dictionary D and obtains a reconstruction that is consistent with the measured sinogram data B. The first two terms in equation (5) optimise the ability of the dictionary to sparsely approximate the image patches and enforces sparsity in the D domain. This equation is subject to a data fidelity constraint in the Radon domain. In the absence of noise, ε should equal zero, otherwise, it should be estimated according to the standard deviation, σ of noise in the measurements. The estimation of ε is critical in order to make this cost function robust to noise. In ET, a rough estimate of the noise can be performed by calculating a power spectrum of a TEM image acquired without a sample. The resulting curve is approximately a Gaussian function that can be characterized by a mean μ and standard deviation σ (Frank, 2010, Chapter 11). ε should be slightly above the estimated noise level. Parameters Section provides more details about tuning this parameter.

This adaptive sparsity-based formula in equation (5) is potentially capable of reconstructing tomograms using only the under-sampled sinogram measurements and prior knowledge about the noise level. By regularisation using a patch-based dictionary, missing wedge artifacts and noise can be further suppressed, without introducing new artifacts, compared to other transforms that are based on the relationship of neighbouring pixels. Since patches are extracted with overlap for training the dictionary, the DLET algorithm benefits from an additional averaging operator that can be effective for de-noising and reducing reconstruction artifacts. Also, the patch size can be altered depending on the size of the features in the reconstructed image, and this can be very effective in allowing this technique to perform well in separating strong noise from weak structures.

Proposed Algorithm

The image reconstruction optimisation problem is solved using an alternating minimisation to optimise x, D and α. We follow three stages to solve the problem in equation (5) using the proposed DLET algorithm.

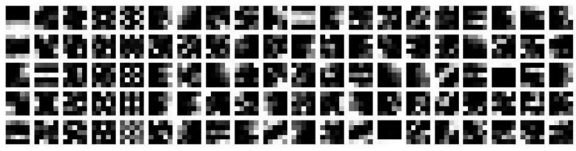

Fig 5. An adaptively learned dictionary consisting of (K = 100) atoms of 5×5 patches, normalized to the range of [-1,1].

Algorithm III.1: DLET(B, n, K, c, σ)

Data: B - Projections (sinogram)

Initialize: $t \leftarrow 1$, $\mathrm{x}^t \leftarrow \mathrm{R}^{-1}B$, $\lambda \leftarrow \frac{c}{\sigma}$

while Stopping Criteria is not satisfied **do**

(1) Construct Training set from x^t

(2) Learn Dictionary D using KSVD

(3) Find sparse coefficients α using BPDN

(4) Compute intermediate x^* from Dα

(5) Compute sinogram B^* of x^* using $\mathrm{R}_{\theta=(0:180)}$

(6) $\mathrm{B}^{\varphi} \leftarrow \begin{cases} \lambda B + \frac{B^*}{\lambda+1} & \forall\ \theta \text{ sampled in B.} \\ \mathrm{B}^* & Otherwise. \end{cases}$

(7) Construct a new Training set from B^{φ}

(8) Apply steps (2-4) using new Training set

(9) Compute intermediate B^t

(10) $\mathrm{x}^{t+1} \leftarrow \mathrm{R}^{-1}B^t$

return (x)−Reconstructed Image

Fig 6. DLET algorithm to reconstruct ET images from undersampled tilt serises.

1. On-line Dictionary Learning: In the first stage, the reconstructed image x is initialised using the weighted back projection (WBP) method in order to reduce data discrepancies. Then, x is fixed and the dictionary D and the sparse representation α are updated by solving the following subproblem:

$$D_t \triangleq \min_D \frac{1}{S} \sum_i^s \frac{1}{2} \|M_i x - D\alpha_i\|_2^2 \, s.t. \, \|\alpha_i\|_1 \leq n_{nz} \quad (6)$$

Typically, this optimisation problem can be solved efficiently using the K-SVD algorithm (Aharon *et al.*, 2006a) to learn the dictionary D dynamically from x during the reconstruction process. We used the fast online dictionary-learning algorithm (Mairal *et al.*, 2010) to train D. To avoid any scaling problems that may arise, atoms are produced with unit norm. After learning the dictionary, a sparse coding step is executed.

2. In the second stage, keeping x and D fixed, α is updated according to the cost function:

$$\alpha_t \triangleq \min_\alpha \|\alpha_i\|_1 \, s.t. \, \|M_i x - D\alpha_i\|_2^2 \leq \varepsilon \quad (7)$$

This formulation is called the basis pursuit denoising (BPDN) problem. Solving this is very efficient especially if there is inherent noise in the measured signal (van Den Berg and Friedlander, 2008; Tropp and Wright, 2010). For the ET reconstruction, we used a Poisson model of shot noise as in (Williams and Carter, 2009; Scott *et al.*, 2012). Sparse coding is applied to all image patches $M_i x$ in order to determine the sparse coefficient α.

3. Finally, the solution x is updated by reassembling overlapping patches using the transpose of M as $\sum M_i^T D\alpha_i$ and constructing an intermediate image x^*. In this step, the intensity value of a pixel in x^* is

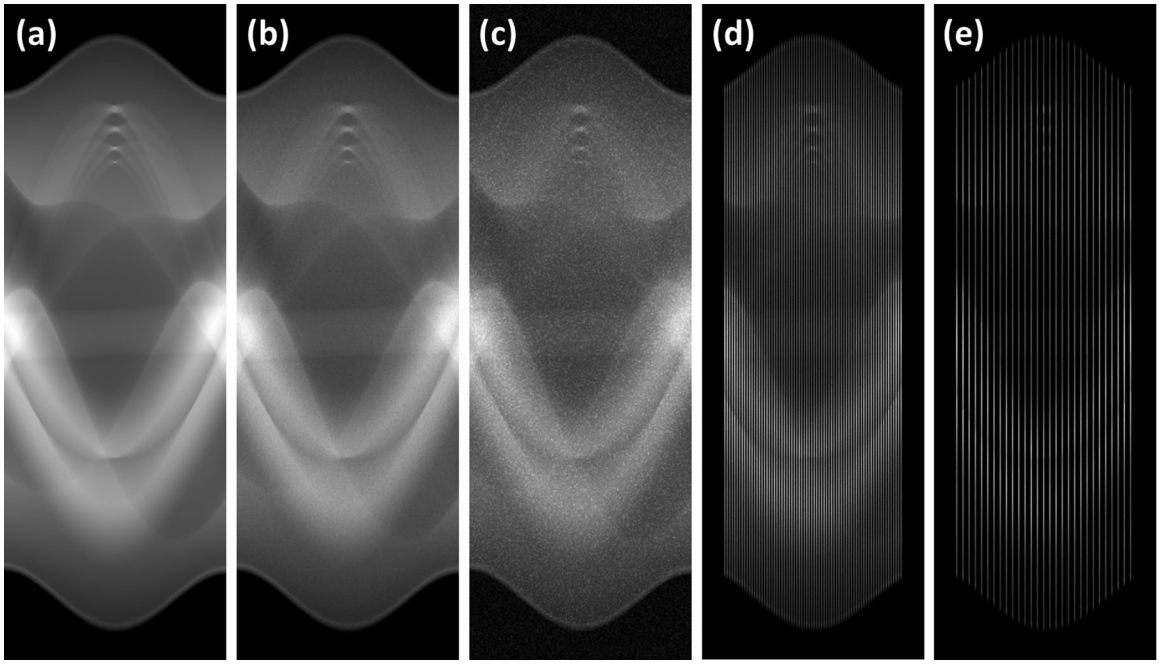

Fig 7. Simulated Sinogram data from the phantom shown in (Fig.7a). (a) Noiseless fully sampled true sinogram (512 pixel × 180°). (b) and (c) noisy fully sampled sinogram with SNR of 52dB and 15dB, respectively. (d) and (e) are noiseless undersampled sinogram with 70 and 29 radial lines, respectively (leaving a 20° missing wedge of un-sampled information).

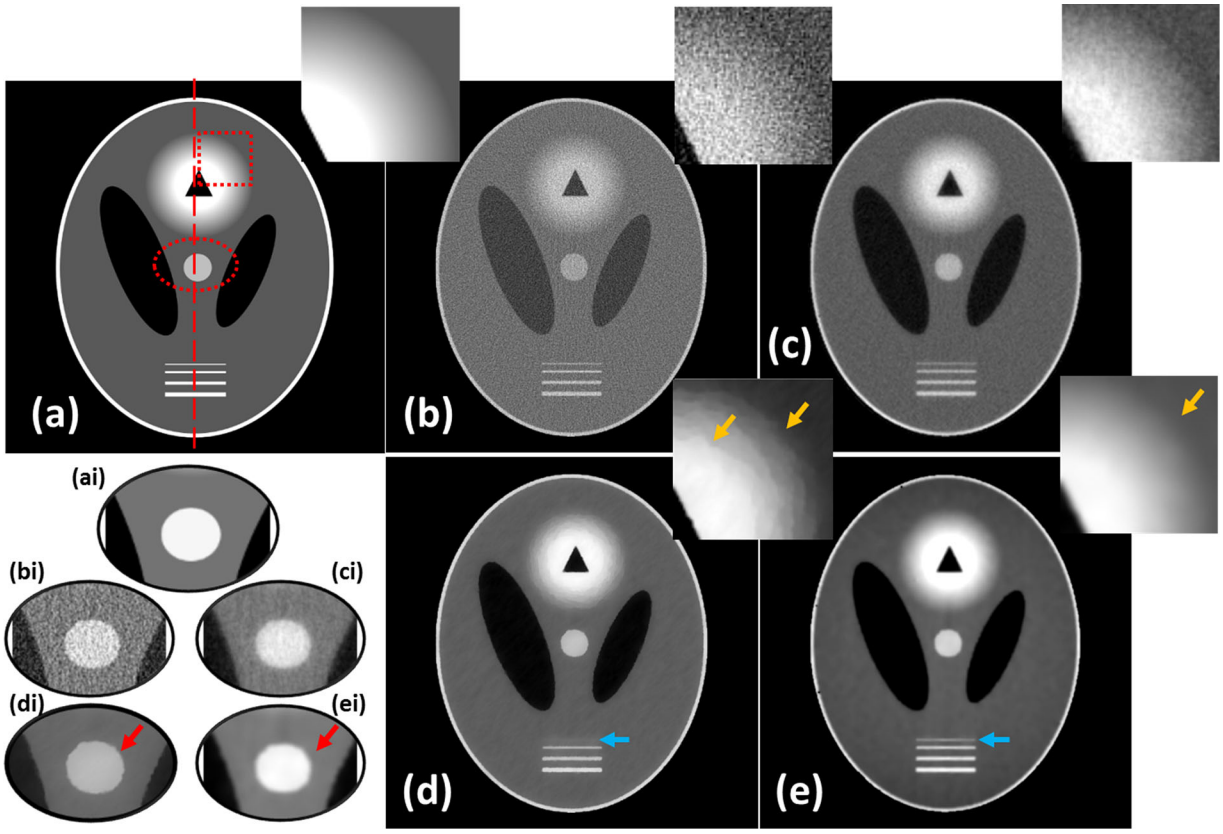

Fig 8. Comparison of different reconstruction algorithms using the noisy fully sampled sinogram in (Fig.6b). (a) Ground truth, compared to reconstruction using (b) WBP, (c) SIRT, (d) CSTV with $\lambda_{TV} = 7$ and $\lambda_{l1} = 0.1$, f) DLET with $\varepsilon = 1 \times 10^{-7}$. (ai-ei) are zoomed regions extracted from the phantom and reconstructions at the area indicated by oval-dotted region in (a). For visualisation purposes, the image contrast of zoomed areas is enhanced.

computed by averaging all the corresponding patches that cover it. The image x^* is then be projected to the fully sampled Radon domain using $R_\theta x^*$ to get B^*. Both the new B^* and the measured under-sampled B are combined to get B^φ, and another iteration of Dictionary learning and Sparse Coding is applied to B^φ to computer B^t. Finally, the new x^{t+1} is formed by $R^{-1}B^t$ as detailed in Figure 6. The above procedures are iterated until a stopping criterion is reached.

Parameters

The proposed algorithm requires user input for a few parameters:

- Discrepancy level (ε): controls the allowed tolerance between reconstructed and measured data. Typically, this is mainly affected by the noise standard deviation σ in the measured data and can be computed as: $\varepsilon = (C\sigma)^2$, where C is a scaling factor.
- Patch size (n): controls the dictionary efficiency in learning features of an image. A large patch size produces more dictionary atoms, which increases the computing cost of the algorithm.
- Number of atoms (k): controls degree of over-completeness (redundancy) of the dictionary which improves the sparsity of representation. The number of dictionary atoms should be larger than patch size to ensure good redundancy. In Elad and Aharon (2006) it was shown that a redundancy of $K = 4n$ would be sufficient for image denoising problems.

We tested the sensitivity of the DLET algorithm to these parameters in section IV.e.

Objective Function Convergence

The non-negative cost function in equation (5) is solved by alternating between dictionary learning (sub-cost function equation (6)), sparse coding (sub-cost function equation (7)), and updating the solution x. Each of these terms in the cost function decreases monotonically on iteration, meaning that the objective function in equation (5) also decreases monotonically using the proposed algorithm. However, this monotonic decrement does not guarantee the convergence of the reconstruction process. The convergence of DLET is difficult to prove, and beyond the scope of this paper.

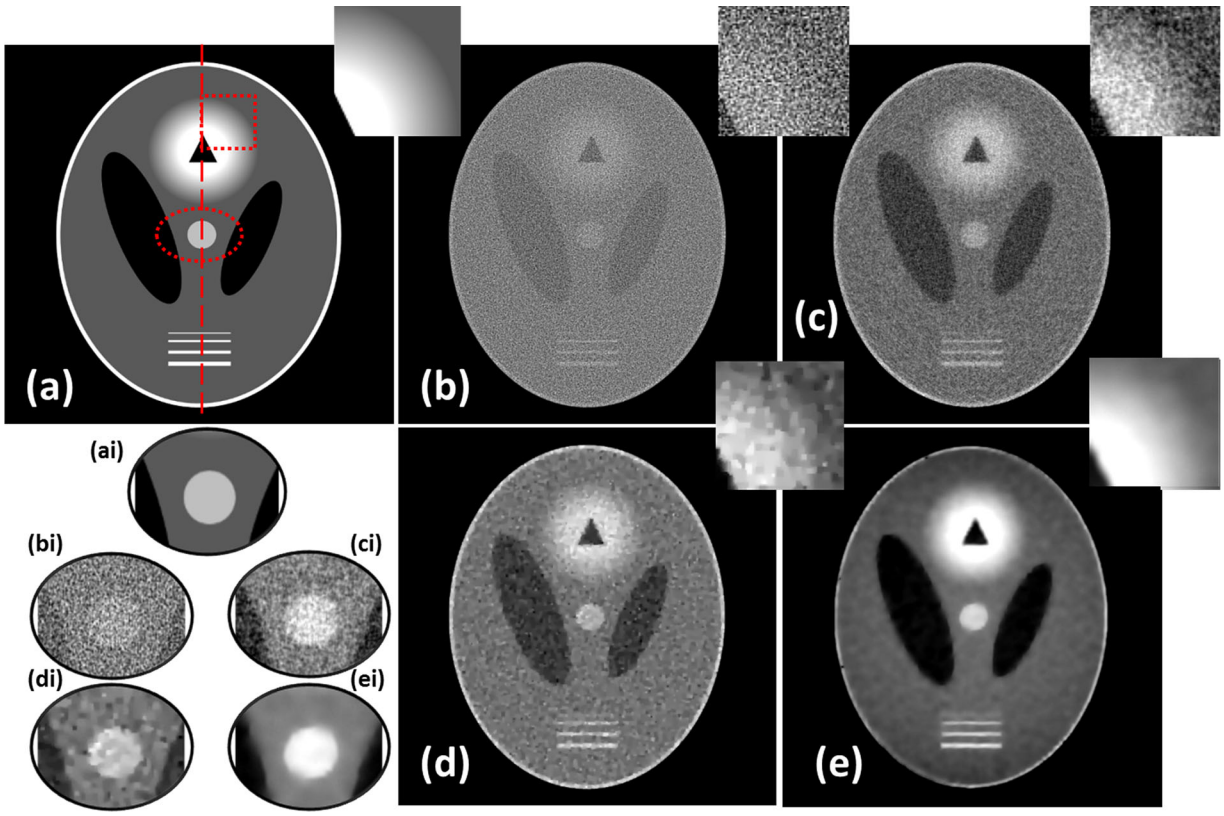

Fig 9. Comparison of different reconstruction algorithms using the noisy fully sampled sinogram in (Fig.6c). (a) Ground truth, compared to reconstruction using (b) WBP, (c) SIRT, (d) CSTV with $\lambda_{TV} = 12$ and $\lambda_{l1} = 0.1$, and (f) DLET with $\varepsilon = 2 \times 10^{-6}$ are zoomed regions extracted from the phantom and reconstructions at the area indicated by oval-dotted region in (a). For visualisation purposes, the image contrast of zoomed areas is enhanced.

However, our empirical studies suggest that the proposed algorithm converges in a well-behaved manner.

Termination Criterion and Computational Complexity

The algorithm process is repeated until a plateau is reached such that successive iterations no longer produce better results. For the simulation study, the algorithm was terminated when a fixed number of iterations is reached. This number is chosen to fall after the convergence plateau. For practical implementation, other common terminating conditions may be included, such as allocated budget (computation time/memory) reached, manual inspection, or combinations of the above.

In this work, both dictionary learning and sparse coding are achieved using the SPAM toolbox (Mairal *et al.*, 2010), which is the state-of-the-art approach at the current time and significantly faster than other sparse optimisation methods.

Numerical Simulation

An enhanced ET reconstruction algorithm should be able to avoid artifacts typically seen in conventional ET reconstructions. Artifacts arise mainly for three reasons: missing wedge, radial undersampling, and noise in the measurements. The first two problems cause aliasing in the image domain in the form of streak, blurring, and elongation artifacts. The noise in TEM microscope images depends on the experimental conditions; however, two types of noise are permanently present: the quantum noise of the electron beam (shot noise) and the noise originated from the image recording system which is usually characterised by a modulation transfer function (MTF) and a detective quantum efficiency (DQE) (Frank, 2010, Chapter 11). The propagation of noise from the sinogram domain into the reconstructed volume domain is a complicated process.

To evaluate the performance of the proposed reconstruction algorithm, two simulation case studies were performed. In the first case, a modified version of the well-known Shepp–Logan phantom (Fig. 8(a)) is used. For the second case, a more challenging phantom is used with features that are hard to reconstruct and which are not sparse in the gradient domain as shown in Figure 13. For comparison, we performed the reconstruction using the following methods: WBP, SIRT, a CSTV-based method, and the proposed DLET approach, using a range of different under-sampling factors and noise levels.

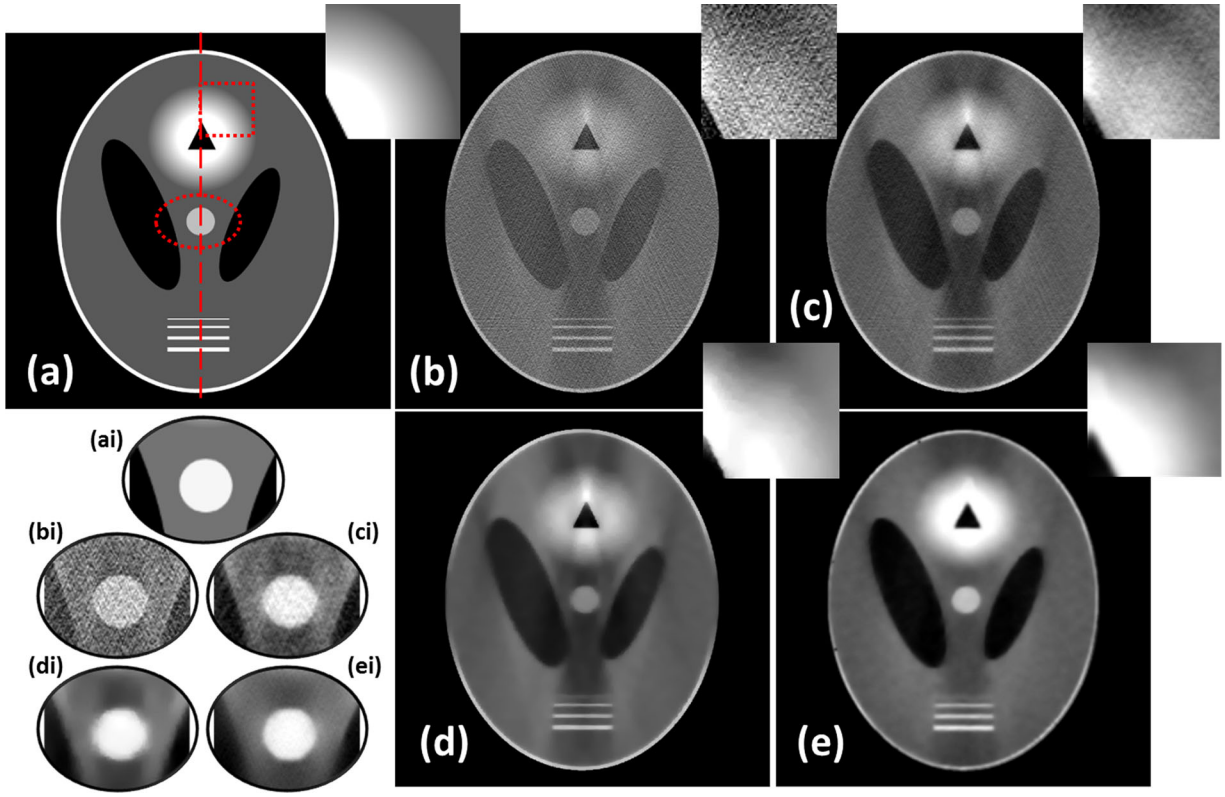

Fig 10. Comparison of different reconstruction algorithms using the noiseless undersampled sinogram in (Fig.6d) with 70 radial lines. (a) Ground truth, compared to reconstruction using (b) WBP, (c) SIRT, (d) CSTV and (e) DLET with $\varepsilon = 45 \times 10^{-7}$. (ai-ei) are zoomed regions extracted from the phantom and reconstructions at the area indicated by oval-dotted region in (a). For visualisation purposes, the image contrast of zoomed areas is enhanced.

Image Quality Metrics

With the phantom images as the ground truth, all the reconstructed cases were assessed in terms of two commonly used metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural SIMilarity (SSIM) index. The PSNR was obtained as the ratio between the signal's maximum power (peak reference intensity) MAX_I to the power of the signal's noise (root mean square of reconstruction error) $MSE(I_o - I_R)$ as in equation (8).

$$\mathrm{PSNR}(I_o, I_R) = 10 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE(I_o, I_R)}}\right) \quad (8)$$

The PSNR is measured in decibels (dB) and the higher the PSNR value, the better the quality of the reconstruction.

Although the PSNR is a simple mathematical metric that is commonly used as a distortion metric, it often fails to correlate closely with perceived image quality (Chandler and Hemami, 2007). Consequently, we have chosen to use an additional, more advanced metric, the SSIM index.

The SSIM index is shown to be consistent with visual perception (Wang *et al.*, 2004). The calculation of the SSIM index for the two images ζ and ϱ to be compared begins with computing three similarities: luminance (ζ, ϱ), contrast $c(\zeta, \varrho)$, and structure $s(\zeta, \varrho)$ similarity. The SSIM is defined as:

$$\begin{aligned}\mathrm{SSIM}(\zeta, \varrho) &= l(\zeta, \varrho) \cdot c(\zeta, \varrho) \cdot s(\zeta, \varrho) \\ &= \left(\frac{2\mu_\zeta \mu_\varrho + C_1}{\mu_\zeta^2 + \mu_\varrho^2 + C_1}\right)^{\alpha 1} \left(\frac{2\sigma_\zeta \sigma_\varrho + C_2}{\sigma_\zeta^2 + \sigma_\varrho^2 + C_2}\right) \\ &\quad \beta 1 \left(\frac{2\sigma_{\zeta\varrho} + C_3}{\sigma_\zeta^1 \sigma_\varrho^1 + C_3}\right)^{\gamma 1} \end{aligned} \quad (9)$$

where μ_ζ and μ_ϱ are local means, σ_ζ and σ_ϱ are local standard deviations, and $\sigma_{\zeta\varrho}$ is the cross-correlation after subtracting corresponding means. C_1, C_2, C_3 are stabilisers and $\alpha 1 > 0, \beta 1 > 0, \gamma 1 > 0$ are parameters used to adjust the relative importance of the three components. The maximum value of the SSIM index is 1, which indicates perfect structural similarity. In this work, we used the default values for the parameters in equation (9) as in Wang *et al.* (2004) with $C_1 = (0.01L)^2, C_2 = (0.03L)^2, C_3 = \frac{C_2}{2}, L = 255$ and $\alpha 1 = \beta 1 = \gamma 1 = 1$.

For further information on image quality metrics, the interested reader is referred to the review of Sayood (2002).

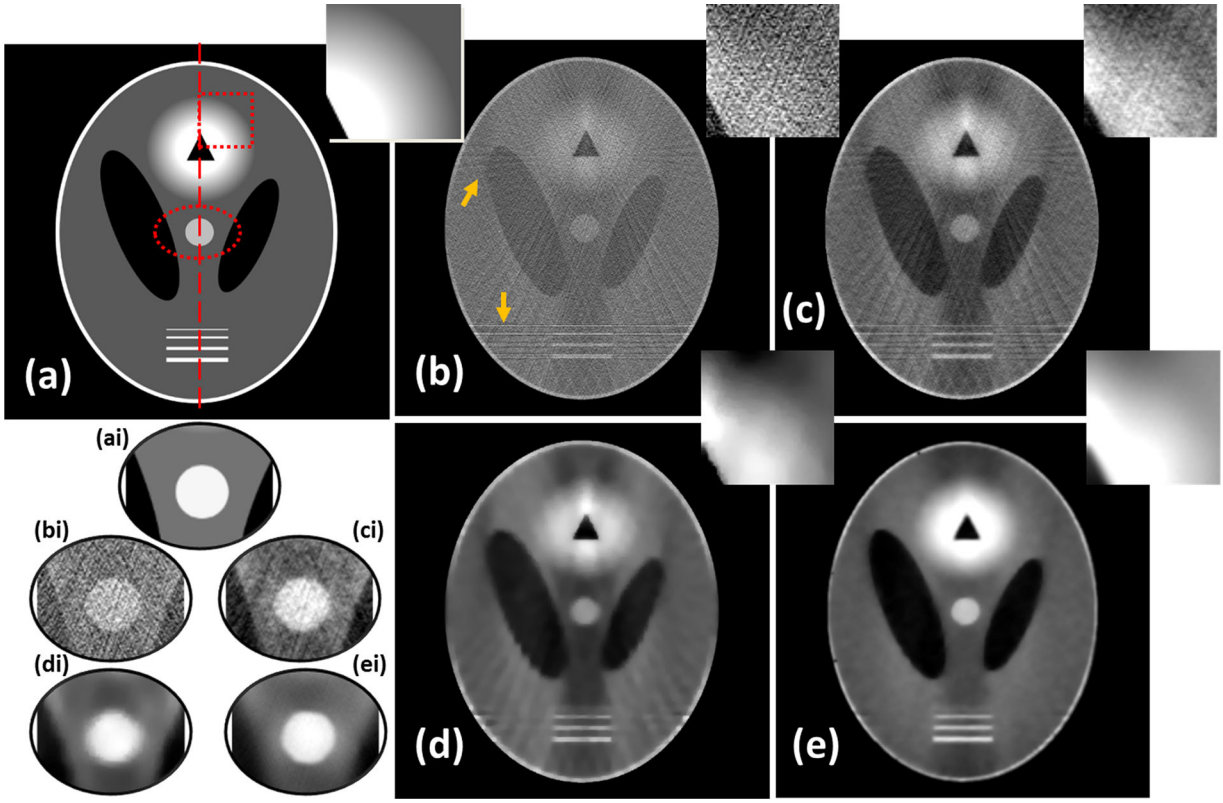

Fig 11. Comparison of different reconstruction algorithms using the noiseless undersampled sinogram in (Fig.6e) with 28 radial lines. (a) Ground truth, compared to reconstruction using (b) WBP, (c) SIRT, (d) CSTV and (e) DLET with $\varepsilon = 45 \times 10^{-7}$. (ai-ei) are zoomed regions extracted from the phantom and reconstructions at the area indicated by oval-dotted region in (a). For visualisation purposes, the image contrast of zoomed areas is enhanced.

Case Study 1: The Modified SheppLogan Phantom

In this section, we compare the proposed DLET technique with conventional ET methods, namely, WBP, SIRT, and the CS-based total variation (CSTV) method. We used the simulated phantom shown in (Fig. 8(a)). This phantom was edited to include a bright circle region in the top part with gradual intensity variation in the background. Such intensity variations can simulate realistic materials in a normal STEM experiment such as solar cells (van Bavel and Loos, 2010). Also, horizontal lines with different thicknesses were added to verify the degree of detail that each algorithm can preserve. Two simulation setups were performed. The first one aimed to test the case of a noisy fully sampled sinogram as in Figure 7(b and c) while the second test a noiseless under-sampled sinogram with a missing wedge as in Figure 7(d and e). In order to avoid committing an inverse crime, which happens when the data is inappropriately simulated (Kaipio and Erkki, 2006, chapter 1), the tilt series were generated using the parallel projection discrete Radon transform, while the reconstruction was coded using Fourier-based methods so that a different system matrix is used for creating projections than in the reconstruction methods. Furthermore, the tilt series was misaligned by randomly shifting each projection by a maximum of $\pm 0.5^{\circ}$ to account for the alignment imperfections in the experimental ET data. By doing so, critical errors can be avoided when these methods are applied to real ET data.

The input data in each case was then prepared by taking the 1D Fourier transform of each projection in a tilt series and sampling it to the corresponding radial line in the 2D Fourier domain using a sampling mask. The WBP reconstruction was performed with a ramp filter in the frequency domain to produce an image in the spatial domain. This image was then used to initialise the SIRT, CSTV, and DLET algorithms. SIRT was performed using 32 iterations. This number was selected after investigating which number of iterations gives the maximum quality in terms of PSNR and SSIM metrics by running SIRT for 100 iterations and recording the quality metric values for each iteration. CSTV reconstruction was performed using the method provided by Lustig *et al.* (2007b) seeking sparsity in gradient and image domain with default regularisation weighting of $\lambda_{TV} = 7$ and $\lambda_{l1} = 0.1$, respectively. This provides a balance between the loss of fine details and the elimination of ghosting artifacts. With CSTV, 150 conjugate gradient iterations were performed with a re-initialisation every 50 iterations in order to decrease the likelihood of falling into local minima. DLET was

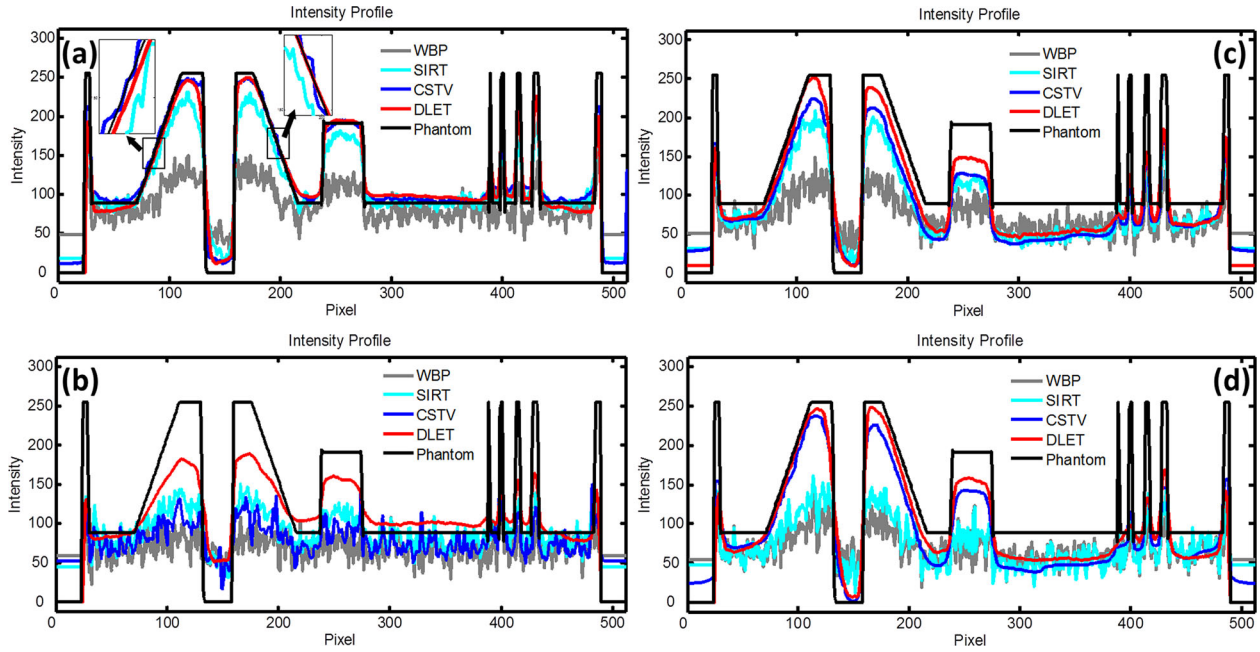

Fig 12. Intensity line profile comparison along the vertical dashed line in a) Fig. 7(a), b) Fig. 8(a), c) Fig. 9(a) and d) Fig. 10(a). Each figure is a comparison between reference phantom image with reconstruction from WBP, SIRT, CSTV, and DLET for different test cases.

run for 20 iterations with the parameters chosen as $n = 64, K = 4n$ and$T_0 = 0.5n$. The dictionary learning stage using K-SVD was performed for 20 iterations, with 10^3 overlapping patches extracted from the intermediate image as a training set and a target sparsity of T_0. The boundary condition is assumed to be reflective in the DLET training set. All implementations were executed on Matlab v7.12 (R2011a) installed on a 64-bit Windows 7 operating system with an Intel Core i5 processor running at 3.10 GHz with 24 GB RAM and a NVIDIA GeForce GTX 460 with 336 cores.

Noisy Full-Sampled Setup

To compare ET reconstruction methods with noisy data, the Phantom in (Fig. 8(a)) of 512×512 pixels, was projected into the sinogram domain (Fig. 7(a)) between $\pm 90^{\circ}$ with 1° tilt increment. This was then corrupted by applying (i) poisson noise to simulate the shot noise and (ii) Gaussian noise with low (Fig. 7(b)) and high (Fig. 7(c)) standard deviation σ to simulate electronic noise in the amplifier system. Figures 8(b) and 9(b) shows WBP reconstructed images with obvious artifacts

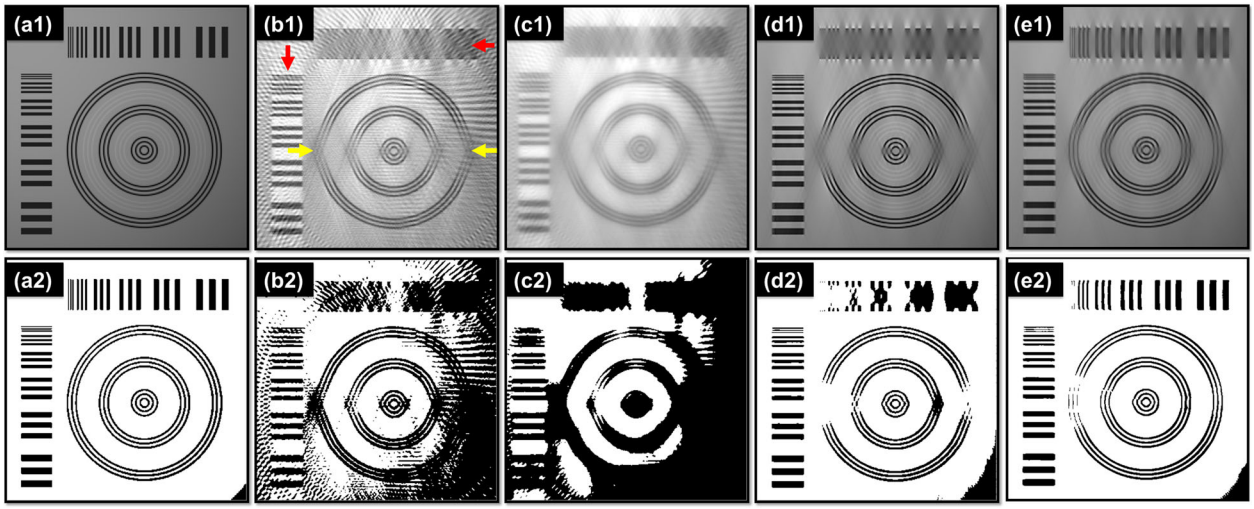

Fig 13. Visual assessment of the missing wedge artifact. The top row is an extract from the lower-right quarter of each of the top row images in Fig. 12. a1) Reference phantom extracted from image in (Fig.11 b1) WBP image in (Fig.12a1), c1) SIRT image in (Fig.12b1), d1) DLET image in (Fig.12c1) and e1) DLET using global dictionary. The bottom row shows the corresponding thresholded images obtained after applying automated thresholding using Otsu's method.

and noise. Based on the visual assessment, WBP with a Ramp filter cannot successfully remove artifacts in reconstructions. Figures 8(c) and 9(c) shows the results of the SIRT algorithm with improved reconstruction and contrast (SSIM = 0.3478, 0.1555) compared to (0.1748, 0.0886) for WBP. (Fig. 8d), CSTV reconstruction shows an effective contrast and noise suppression in homogeneous regions. However, this improvement comes at the cost of losing the fine details of structures and of introducing staircase artifacts in the region with gradual changes of intensity (blue and orange arrows in zoomed images in (Fig. 8); these staircase artefacts cannot be suppressed by increasing the λ_{TV} parameter values. In Figure 9(d), CSTV was tuned to de-noise by increasing λ_{TV}; however, it was found that increasing λ_{TV} above 12 caused the loss of fine details.

Figures 8(e) and 9(e) show results obtained after applying the DLET algorithm, and these are promising showing both noise suppression and the preservation of fine structures. The degradation of visual quality with respect to the phantom image was noticeably lower than for other algorithms when applied to the sinogram with the higher noise level, with a much lower percentage decrease the of SSIM index (Δ SSIM $= -5.01\%$) compared to CSTV (-63.39%), SIRT (-55.29%), and WBP (-49.31%).

Figure 12(a) and (b) show the intensity line profile comparison along the vertical dashed line in Figures 8(a) and 9(a). It can be seen that reconstruction using DLET and CSTV leads to images with a better match to the phantom than the WBP and SIRT results. It is also noticeable that both the CSTV and DLET reconstruction methods can provide a relatively accurate recovery of the homogenous area (see the red arrows in zoomed area of Fig. 8). In addition, DLET is more robust in cases with high noise (Fig. 9(e)) with better preservation of fine details (see blue arrows in Fig. 8) and regions with gradual intensity variation (see orange arrows in Fig. 8). Table I lists the PSNR and SSIM values computed with respect to the reference phantom images in (Fig. 8(a)).

Noiseless Under-Sampled Setup

To verify how each method behaves in the case of undersampling, projections (columns) were removed from the sinogram in (Fig. 7(a)) to simulate under-sampling between $\pm 70^{\circ}$ with tilt increments of 2°(Fig. 7 (d)), and 5° (Fig. 7(e)).

As a consequence of such under-sampling, aliasing effects can be seen in both cases as streak artifacts. Also, the missing 20° wedge, located on both sides of sinograms, causes another form of aliasing (i.e., elongation artifacts) which increases the object length in the horizontal direction. Such artifacts can be clearly seen in Figures 10 and 11(b and c) which show results obtained by WBP and SIRT methods.

Figures 10 and 11(d) and (e) show that both DLET and CSTV result in reduced under-sampling artifacts. In Figure 10, the DLET reconstruction shows sharper edges with SSIM = 0.8256 compared to 0.4113 of CSTV. In Figure 11, it was harder to reduce artifacts with the increasing under-sampling. However, DLET showed a slightly better visual quality with SSIM = 0.7857 compared to 0.3656, 0.2304, 0.1324 for CSTV, SIRT, and WBP, respectively. Figure 12(c) and (d) shows the intensity line profile comparison along the vertical dashed line in (Fig. 10a) and (Fig. 11a).

Case Study 2: The CS-Phantom

In this section, the performance of the DLET reconstruction algorithm is evaluated using the test phantom (the CS-phantom) proposed by Smith and Welch (2011), which is not a piecewise constant image. The CS-phantom (Fig. 13(a1)) is designed for testing the accuracy of CS solvers and the properties of CS reconstruction artifacts. For comparison, we performed the reconstruction based on WBP, SIRT, and DLET method using simulated dataset. The CSTV-based method was excluded because the CS-phantom is not sparse under a gradient transform, since this violates the requirement for a TV-L1 minimisation. The DLET method was applied using two setups. In the first setup, the dictionary was trained using intermediate (Local) images during the reconstruction process, while in the second setup, a (Global) dictionary was trained using simulated examples that contain features similar to those in the original phantom.

An under-sampled tilt series was modelled with 70 noiseless projections with 2° increments between each projection. The simulation setup described in section (IV-B) was adopted in preparing sinograms for reconstruction. To further assess the visual quality of the reconstructions, the tomograms were thresholded

TABLE I Noise and undersampling evaluation-quality values of modified shepp–logan phantom

Figure	8		9		10		11	
Method/metric	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
WBP	14.40	0.1748	12.19	0.0886	14.61	0.2414	12.64	0.1324
SIRT	20.88	0.3478	14.71	0.1555	16.61	0.2926	15.74	0.2304
CSTV	21.35	0.8552	15.45	0.3131	17.68	0.4113	17.31	0.3656
DLET	21.52	0.8710	17.39	0.8274	18.5941	0.8256	18.24	0.7857

TABLE II Quality metric values of cs-phantom reconstruction inFig. 13

Figure	11(b1-c1)	
Method/metric	PSNR	SSIM
WBP	14.20	0.5651
SIRT	18.73	0.6164
DLET (Local D)	24.97	0.8188
DLET (Global D)	26.57	0.8964

TABLE III Patch size parameter evaluation of DLET — quality values for different runs of DLET

Patch size (n)	PSNR	SSIM	Time (mins)
2	24.34	0.6724	5
4	24.54	0.7647	7
6	24.89	0.8058	15
8	24.97	0.8188	44
10	25.00	0.8224	121

based on the image intensity followed by binarization (i.e., setting pixels above a threshold limit to a value of 1/white and the rest to zero/black). The threshold values were obtained via automated thresholding using Otsu's method (Otsu, '75).

Figure 13 shows the reconstruction using WBP, CSTV, and DLET with two setups. In (Fig. 13(b2)), two types of artifacts can be seen (i) "streaking" artefacts, which is a prominent kind of aliasing caused by the finite and limited angular sampling in ET, and (ii) elongation and blurring of the object boundaries (see horizontal red arrow), which occurs primarily due to the missing wedge of unsampled data. The missing wedge direction runs horizontally between the left and right of the image. The high artefact levels are clear in Figure 13(b1–c1, b2–c2) for both the SIRT and WBP reconstructions, including the well-known artefact of elongation of the reconstructed object in the missing wedge (horizontal) direction. For the DLET reconstruction, Figure 13(d1, d2) and (e1, e2), show further reduction of streaking background artifacts and missing wedge elongation.

The missing wedge direction and object morphology is an important factor that affects the accuracy of the reconstruction. The boundaries of the parallel horizontal lines (indicated by a vertical red arrow) are well reconstructed. This is due to the presence of the significant projection for this shape which is the projection at 0° (horizontal direction) that provides the information about width and distances between each line. However, the horizontal lines (indicated by horizontal red arrow) are difficult to distinguish as the most important projection is in the middle of the missing wedge region (90°). The DLET with global dictionary was able to compensate for most of the missing widget artifacts when trained using similar examples, however, such examples are not usually available in the case of ET. Although the DLET with local dictionary was not able to compensate for the missing wedge of vertical lines, there was a significant reduction of streaking and elongation artifacts compared to both WBP and SIRT.

The quality metric values are listed in Table II. Quantitatively, the WBP algorithm had the worst results. The SIRT result is slightly better than the WBP result (SSIM of 0.6164 and 0.5651 respectively). However, DLET outperformed both with SSIM 0.8188 with respect to the reference image. Also, the PSNR metric confirms this with values of 24.97 dB compared to 18.73 dB for the SIRT and 14.20 dB for WBP. The DLET with Global dictionary, performed slightly better with SSIM of 0.8964 and PSNR of 26.57. Also, background artifacts and false elongation are markedly reduced.

Based on this simulation study, it is clear that the use of an adaptive dictionary-based reconstruction method suppresses artifacts due to under-sampling, streaking and elongation far more effectively than conventional compressed-sensing electron tomography (Saghi *et al.*, 2011; Binev *et al.*, 2012; Goris *et al.*, 2012) or the well-known weighted back-projection algorithm. Additionally, DLET demonstrated a stable and fast convergence for this particular challenging phantom as shown in various figures suggesting that it will be robust for use in real cases. A particularly promising feature of DLET is the relatively competitive reconstructions than existing methods for significantly under-sampled datasets with

TABLE IV Number of atoms (dictionary size) parameter evaluation of DLET—quality values of different runs of DLET from 70 noiseless projections

Dictionary size	PSNR	SSIM	Time
64	24.6	0.80	7
128	24.9	0.81	8
192	24.9	0.81	9
256	24.9	0.81	10

TABLE V The quality of reconstruction was maximised when the discrepancy level (ε) was increased to 0.001 and declined slightly for higher values

ε	PSNR	SSIM	Time
1×10^{-4}	24.60	0.7246	28
1×10^{-3}	24.97	0.8188	16
5×10^{-3}	24.71	0.8002	8
1×10^{-2}	23.87	0.7847	7
1	23.38	0.7257	7
10	23.38	0.6792	6
100	23.38	0.6624	6

large tilt steps, at least up to a certain point at which artifacts start to creep in. Nevertheless, it seems likely that even this limitation could be avoided by using a global dictionary, in order to preserve high spatial frequency structures effectively and to avoid undesirable artifacts due to the high under-sampling. This would make it especially attractive for low-dose electron tomography of beam-sensitive structures, such as biological ultrastructure or polymer solar cells.

Parameter Comparison

To evaluate the sensitivity of the DLET algorithm to parameter settings, the algorithm performance in reconstructing the reference CS-phantom in Smith and Welch (2011), from 70 projections was investigated by altering a single parameter each time whilst fixing the others at their default values. The parameters assessed were: the Dictionary patch size (n), the Number of atoms (K) and the Discrepancy level (ε). The excution time is recorded for each setup.

In Table III, the quality of reconstruction continues to improve as the patch size increases from 2×2 to 10×10 pixels. This is to be expected, as each pixel will be a result of the averaging of patches with a larger area. However, this improvement comes at the cost of a higher computational expense, and a realistic compromise will have to be struck in real application between quality and computational expense.

In Table IV, the performance metrics reached a plateau when the number of dictionary atoms increased to twice the patch size, meaning that the dictionary contains most of the important atoms that can be found in the training data to produce a good sparsity. The training algorithm used here was initiated using the blocks from the discrete cosine transform (DCT) as in Elad and Aharon (2006) and trained on patches extracted from intermediate reconstruction results.

In Table V, the quality of reconstruction was maximised when the discrepancy level (ε) was increased to 0.001 and declined slightly for higher values. This behaviour can be explained as follows: when ε is set at a precision level that is too small, the algorithm includes under-sampling artifacts and noise in the reconstruction during the sparse coding step, which may introduce fake structures. On the other hand, when ε is too large, the quality degrades as the sparse solver does not include enough important patches to properly approximate the solution which causes a loss of resolution.

Figure 14 showes the quality limit for higher under-sampling factors using default DLET settings. Both the PSNR and SSIM valuse of DLET is high even at very large tilt incriments such as 25, indicating good removal of artifacts and noise.

From these results, it can be seen that satisfactory results can be obtained with little tuning of n and K, as the proposed algorithm is not very sensitive to deviation in these parameters; however, the discrepancy level, ε, needs to be accurately estimated.

Three-Dimensional Morphology of Organic Photovoltaic (OPV) Solar Cells

In this section, the DLET is used to study the morphology of a solar cell sample (PTB7:PC BM blends). In Alekseev *et al*. (2015) SIRT reconstruction was used to investigate the shape of $PC_{71}BM$-rich domains using Energy Filtered TEM (EFTEM). SIRT was able to show a nearly ellipsoidal shape of these domains; however, the reconstruction was noisy which made the segmentation a difficult task. This affects the accuracy of quantitive studies. The sample is a blend of polymer (PTB7) and fullerene derivative ($PC_{71}BM$) spin-coated from chlorobenzene with an average particle size of 200 nm in diameter. The tilt series in Alekseev *et al*. (2015) has a low signal-to-noise ratio (SNR) making it an ideal sample to test the denoising capability of DLET. An EFTEM projection from the tilt series is shown in Figure 15.

Before CS can be applied for tomographic reconstruction of organic solar cells using the EFTEM tilt series, two key criteria need to be satisfied: (i) The contrast mechanisms of the microscope using EFTEM need to obey the projection requirement for tomographic reconstruction (Weyland and Midgley, 2003; Frank, 2006), (ii) a sparsifying transformation must exist for

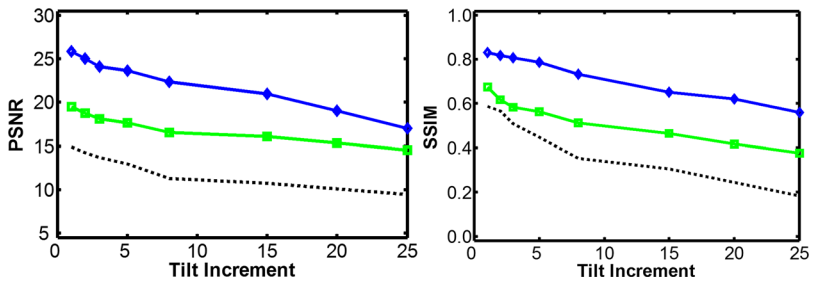

Fig 14. Quality limit for higher under-sampling factors using WBP, SIRT and DLET with higher increment steps between projections.

which the tomogram image is sparse when represented by that transformation.

In order to satisfy the first criterion, the intensity of EFTEM images should be a monotonic function of the projected physical quantity. The PTB7:$PC_{71}BM$ blends sample consists mostly of amorphous material; therefore, the EFTEM contrast is principally determined by the total amount of the specific element mapped in the image (Hofer et al., '95, '97; Weyland and Midgley, 2003) and should not show diffraction contrast. Therefore, the intensity shown in the elemental map of Figure 15 is a projection of the amount of atomic species through the structure of the specimen. As such, EFTEM maps for this particular specimen fulfill the projection requirement for tomographic reconstruction. This also means that the reconstructed volume should contain multiple grey levels, corresponding to a range of different reconstructed intensities of the PTB7:$PC_{71}BM$ sample. It should be also noted that the intensity of the reconstructed grey levels, might not be significantly different due to the noise level in the tilt series arising from the imaging system. For the second criterion, it possible to reconstruct this sample by seeking sparsity in the gradient domain (i.e., optimization using TV) since it is expected for the PC71BM-rich domains to have a uniform composition. However, the boundary between $PC_{71}BM$ domains (carbon-rich) and the PTB7 matrix (sulphur-rich) may not be sharp leading to a range of (non-discrete) grey levels. This makes the reconstruction using TV (i.e., the piecewise constant approximation) inaccurate and may lead to artifacts in reconstruction and lose important details. On the other hand, by using an adaptive data-driven transform via a dictionary learning approach, a sparsifying transform can be tailored to produce a more efficient transform which is crucial for the success of reconstruction using CS. We have performed a simulation study to support the investigation of the experimental data.

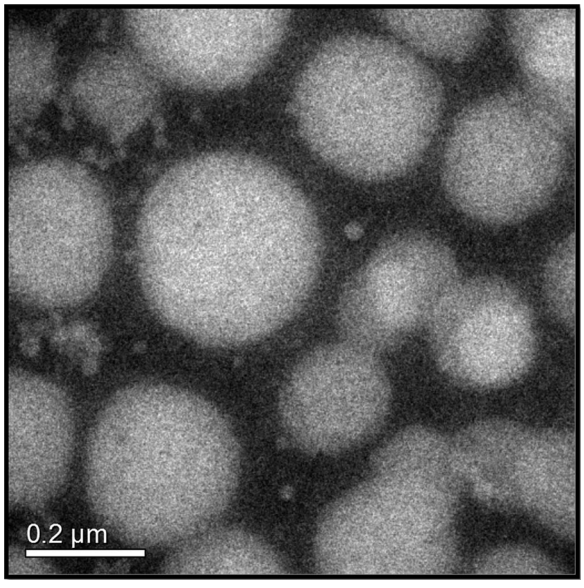

Fig 15. Example EFTEM tomographic tilt-series projection of PTB7:PC BM film. The elemental map of carbon was obtained using the three-window mapping method [Hofer *et al.*, '95; Egerton, 2011] and shows PC71BM-rich domains (large bright regions) embedded in a PTB7-rich matrix.

The 3D Morphology of Polymer Blends for Solar Cells: Simulation Study

To simulate the experimental data, a 3D mesh model of a sphere and two oblate spheroids (see Fig. 16) was designed using CAD software. The sphere was made with a radius of 10 U (top part of Fig. 17(b1)) while the spheroids were made with a fixed value for the equatorial axis ($a = 10$ U in Fig. 16) and a different length for the polar axis ($c = 4$ and 1 units in Fig. 16). Such a model is suitable to simulates the $PC_{71}BM$-rich domains which is believed to have a spherical shape as can be seen in Figure 12 and the cross-section view in Figure 2 in (Alekseev *et al.*, 2015). The simulated mesh model is shown in Figure 17.

This mesh model is then voxelised using the method of Nooruddin and Turk (2003) to generate a binary 3D bitmap with voxel values of logical 1 or 0. The 1 voxels represent the boundary and inside region of the 3D object, while the 0 voxels represent the background region.

The projection tilt series for simulation was generated by applying the discrete Radon transform to the 3D bitmap over a tilt range of $\pm 62^{\circ}$, with an increment of 2° between consecutive projections around the y-axes. Then the tilt series was scaled to match the approximate mean intensities of the $PC_{71}BM$-rich domains evaluated from the experimental EFTEM projections. These were then degraded by the addition of shot and Gaussian noise to get a low SNR dataset that roughly simulates the experimental EFTEM projections. The Gaussian noise was estimated in a way that yields an *SNR* of 7.1, which provides a good

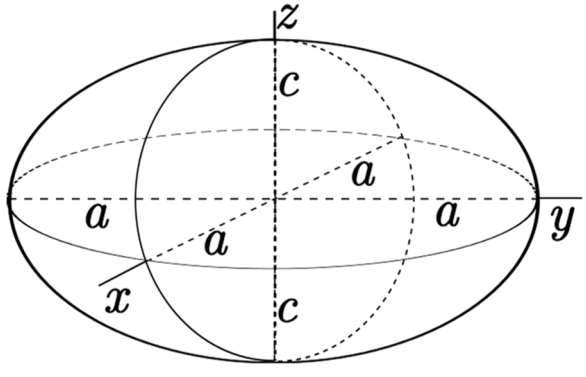

Fig 16. Oblate spheroid: a rotationally symmetric ellipsoid having a polar axis (c) shorter than the diameter of the equatorial axis (a).

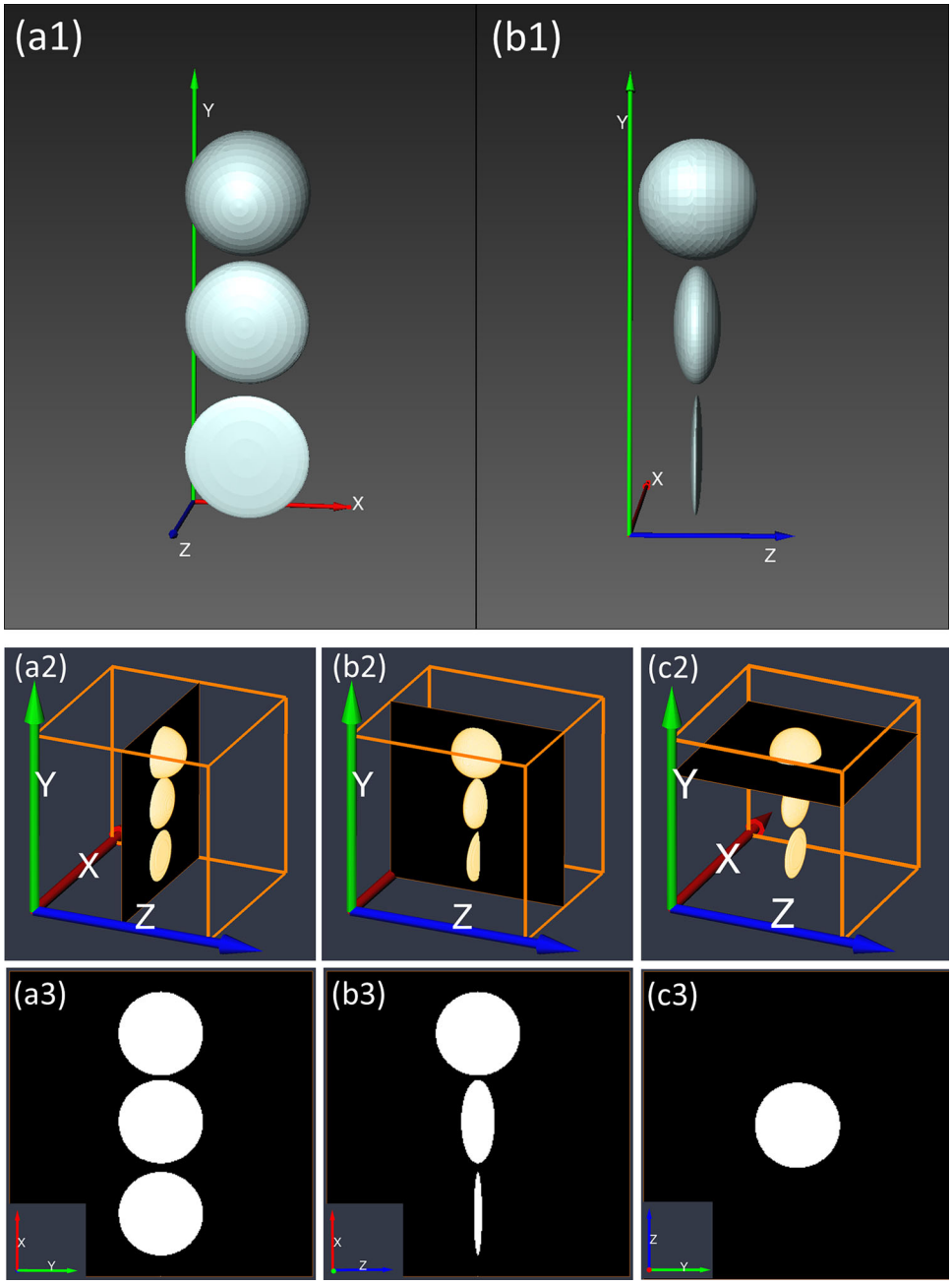

Fig 17. Simulated mesh model for simulating the polymer solar cell sample. The top object is a true sphere of radius, c = 10, the lower two objects are oblate spheroids with different polar axis lengths, c = 4 and 1 units, respectively.

visual match to the noise in the experimental images. Finally, quantisation noise is added. The mapping to quantised intensity was assumed to be linear and is added using equation 10:

$$I_q = \text{floor}\left[2^{Bd}\left[\frac{I_{in}}{I_{\max}}\right]\right] \qquad (10)$$

where I_q is the quantised intensity, Bd is the bit depth, I_{in} is the input intensity from the previous stage and I_{max} is the maximum number of electrons that can be detected by a single pixel of the detector. For this simulation, a quantisation with a bit depth of $Bd =$ gives $2^{16} =$ discrete values of intensity that can be assigned. The parameters for WBP and SIRT were as described in section 3.2. For the CS-TV, finding the appropriate values for λ_{TV},λ_{l1} parameters for a given data set is generally a trial-and-error process which is made more difficult by the subjective nature of deciding what constitutes a 'correct' reconstruction. For this reason, we choose to run the CS-TV algorithm for a range of different values for the parameters between 20–0 for both λ_{TV} and λ_{l1} and choosing the values that maximise the SSIM metric between a 2D XY-slice through the centre in the

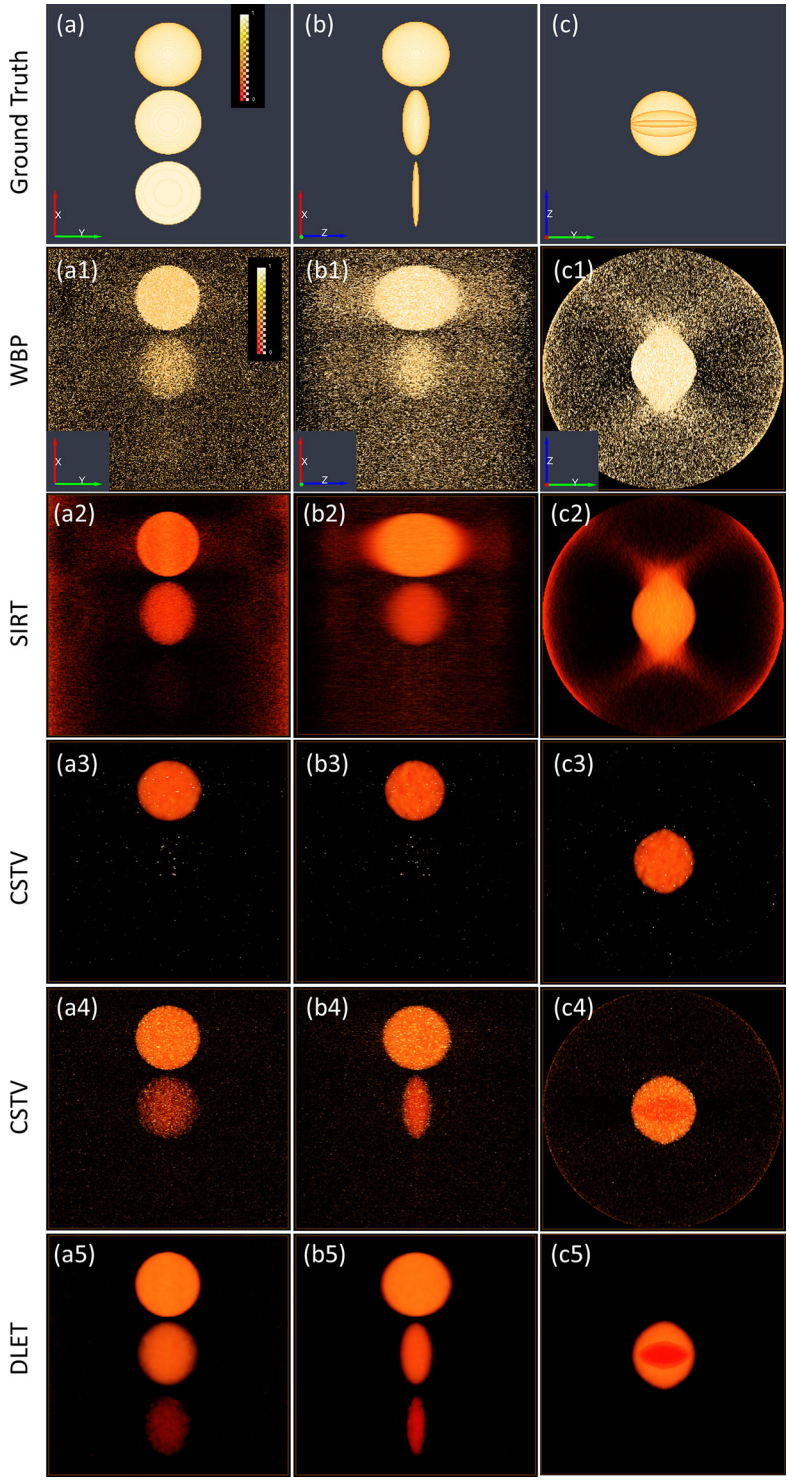

Fig 18. Volume rendering of reconstruction from noisy projections of simulated solar cells- (a1-c1) shows a visualisation WBP (a2-c2) SIRT, (a3-c3) CS-TV with $\lambda_{TV} = 10, 13$, (a4-c4) CS-TV $\lambda_{TV} = 10$, $\lambda_{l1} = 10$, (a5-c5) DLET reconstruction. As before, the a) images are Z-projections, the b) images are Y-projections, and the c) images are X-projections.

simulated 3D map (Fig. 19(a1)) with the corresponding slice in the reconstructed map. The SSIM was shown to be consistent with perceived visual quality as can be seen in Figure 2 in Wang *et al.* (2004). For the DLET method, the dictionary size is set to (K = 256 atoms) for 8×8 patches (n = 64); the dictionary for DL processing is pre-trained from the intermediate image, as a training set, using K-SVD with parameters as follows: 20 iterations are used in dictionary training (Iter = 20); the target sparsity limit T_0 is set to $0.5n$ atoms.

To assess the fidelity of reconstruction quantitatively, we used the normalised Euclidean Distance (NED)

TABLE VI Quality metric values for the polymer solar cell reconstructions

Figure	18			21
Method/metric	SSIM	NED_{img}	NED_{proj}	NED_{proj}
WBP	0.0470	0.74	0.56	(a) 0.42
SIRT	0.0840	0.57	0.31	(b) 0.29
CS-TV	(a3) 0.7454	0.11	0.08	(d) 0.24
	(a4) 0.5755	0.20	0.15	—
DLET	0.8039	0.08	0.07	(f) 0.23

metric in both the image and sinogram domains. In the image domain, it is defined as $\mathrm{NED}_{\mathrm{img}}\|x - x'\gamma\|_2/\|x\|_2$ where x is the ground truth map, x' is the reconstructed image and γ is the scaling constant. In the sinogram domain, the quality metric is: $\mathrm{NED}_{\mathrm{proj}}\|b - \phi b'\|_2\|b\|_2$ where b is the projection data, b' is the decrease Radon transform of x′ and φ is a scaling constant. The scaling factors γ and φ are important for a fair comparison between different algorithms as this will reverse any scaling or negative intensity that might be applied to the simulated data by different numerical implementations used in this article. The values for γ and φ are obtained in a way that minimises the Euclidean distance metric, which is an unconstrained nonlinear optimisation that can be solved using MATLAB.

Figure 18(a1–c1) shows a volume rendering from the WBP reconstruction. This clearly suffers from blurring, streaking, and missing wedge artifacts, which can be also viewed in the noisy orthoslices in Figure 19(a1–c1). The SIRT results in Figure 18(a2–c2) show higher SNR and contrast than of the WBP. SIRT was also able, to some degree, to retrieve the challenging third spheroid that is lost in WBP as indicated by the arrow in (Fig. 19(a2)). In the CS-TV case, the parameters that maximise the SSIM between the reconstructed and ground truth maps were $\lambda_{TV} = 10$, $\lambda_{l1} = 13$ as shown in Figure 18(a3–c3). It can be seen that the CS-TV was able to reduce the elongation artifact and noise, however, this came at the cost of losing both of the spheroids. This problem can be ameliorated by decreasing the λ_{l1} contribution to 10 as can be seen in Figure 18(a4–c4). CS-TV was able, in this case, to retrieve the middle spheroid, however, this comes at the cost of increasing noise in the reconstruction, as can be seen in the orthoslices in Figure 19(a4–c4). It should be also noted that CS-TV reconstruction was not able to retrieve the second spheroid, even with a lower value of λ_{l1} as the noise remains dominant.

Figure 18(a5–c5) shows the DLET results, and these clearly show reduced noise and minimal missing wedge artifacts. The error tolerance parameter ε is set to 2.5×10^5. It can be clearly seen that all of the spheroids were recovered with good contrast. As for the challenging second spheroid, despite the weak signal, it can be seen that this was much better reconstructed by DLET than by any other algorithm. Such an object is challenging to recover as its intensity is very near to the noise level in the sinogram domain, making it hard to distinguish from the noise. The DLET performed better due to the patch averaging step, where each pixel is an average of the image patches that cover the surrounding area. Also, since the other objects contain similar features, this enhances the sparsifying capability of the learned dictionary. This also enables the use of features from other objects with better contrast to recover objects with noisy pixels in a way that is consistent with the measurements. Table VI shows the quantitative quality measurements for each of the reconstructions in this section.

The 3D Morphology of Polymer Blends for Solar Cells: Experimental Results

As stated above, in our previous publication (Alekseev *et al*., 2015), we used ET to study the 3D morphology of PTB7:PC_{71}BM blends, but only using SIRT. Here, we have applied CS–TV and DLET to an EFTEM tilt-series of PTB7:PC_{71}BM blends (Fig. 15), looking for a reconstruction of a higher quality than can be provided by SIRT and WBP. For this purpose, we used a tilt series that was acquired for a cross-section lamella of the polymer blend with thickness (∼130 nm), making it electron transparent. The lamella was prepared using a focused ion beam (FIB) lift-out technique. Such thinning is also important to avoid shadowing at higher tilts. The electron tomography dataset was acquired on an FEI Tecnai T20 TEM operated at 200 kV and equipped with a Gatan GIF2000 Imaging Filter. The energy-filtered transmission electron microscopy (EFTEM) imaging mode was used for energy-selected images for the C-K edge to highlight the PC_{71}BM domains. A tilt series was acquired over a tilt range of ±62°, with an increment of 2° between consecutive projections. After obtaining the elemental map tilt series, an automated spatial drift correction (alignment) for the EFTEM series was performed with the Statistically Determined Spatial Drift algorithm (Schaffer *et al*., 2004) using the SDSD plug-in for DigitalMicrographTM (DM). The visualization of all the reconstructions was done using the Amira 6.0 software package from FEI Visualization Sciences Group. More details about imaging conditions can be found in Alekseev *et al*. (2015).

Figure 20 shows an XY-orthoslice taken at the centre of reconstructed volume using different methods. The WBP and SIRT slices show noisy results in Figure 20(a and b) and missing wedge artifacts in the XZ-orthoslices in Figure 20(ai and bi) which are very difficult to segment. Figure 20(c and d) shows a reconstruction from CS-TV with different values for λ_{TV} =10, 20, respectively. The values for λ_{l1} in both slices was 10 and not increased in order not to lose features from the reconstruction. It can be seen that CS-TV in slice (c),

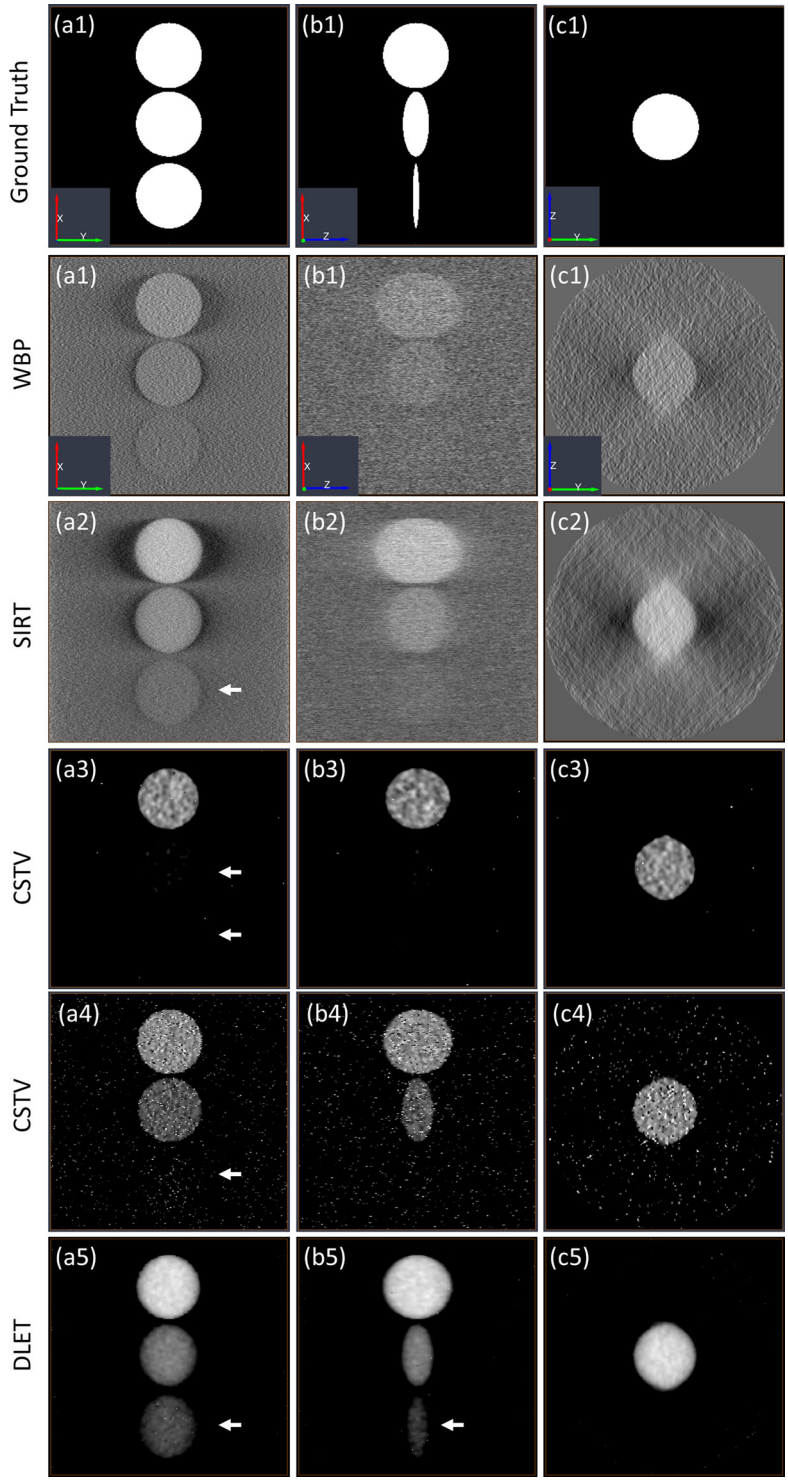

Fig 19. Orthognal Slices through the reconstructed volume in Fig. 19. (a1-c1) shows an orthoslice of WBP reconstruction (a2-c2) SIRT, (a3-c3) CS-TV with $\lambda_{TV} = 10$, $\lambda_{l1} = 13$, (a4-c4) CS-TV $\lambda_{TV} = 10$, $\lambda_{l1} = 10$, (a5-c5) DLET reconstruction. As before, the a) images are sliced perpendicular to Z, the b) images are slices perpendicular to Y, and the c) images are sliced perpendicular to X.

using parameter values similar to those used in the simulation in (Fig. 19(a4)), has a limited ability to reduce the noise. When increasing the λ_{TV} from 10 to 20 in slice (d), the reconstruction noise was further suppressed, however, this came at the expense of introducing staircase artifacts (as indicated by the arrows in the zoomed area). The DLET slices in Figure 20(e and f) show reconstruction with higher fidelity for this noisy and reduced dataset. The results are improved when increasing the error tolerance parameter ε from 2.5×10^5 to 1.3×10^4 in slice (f), which makes it very easy to segment using (semi-) automated methods. In

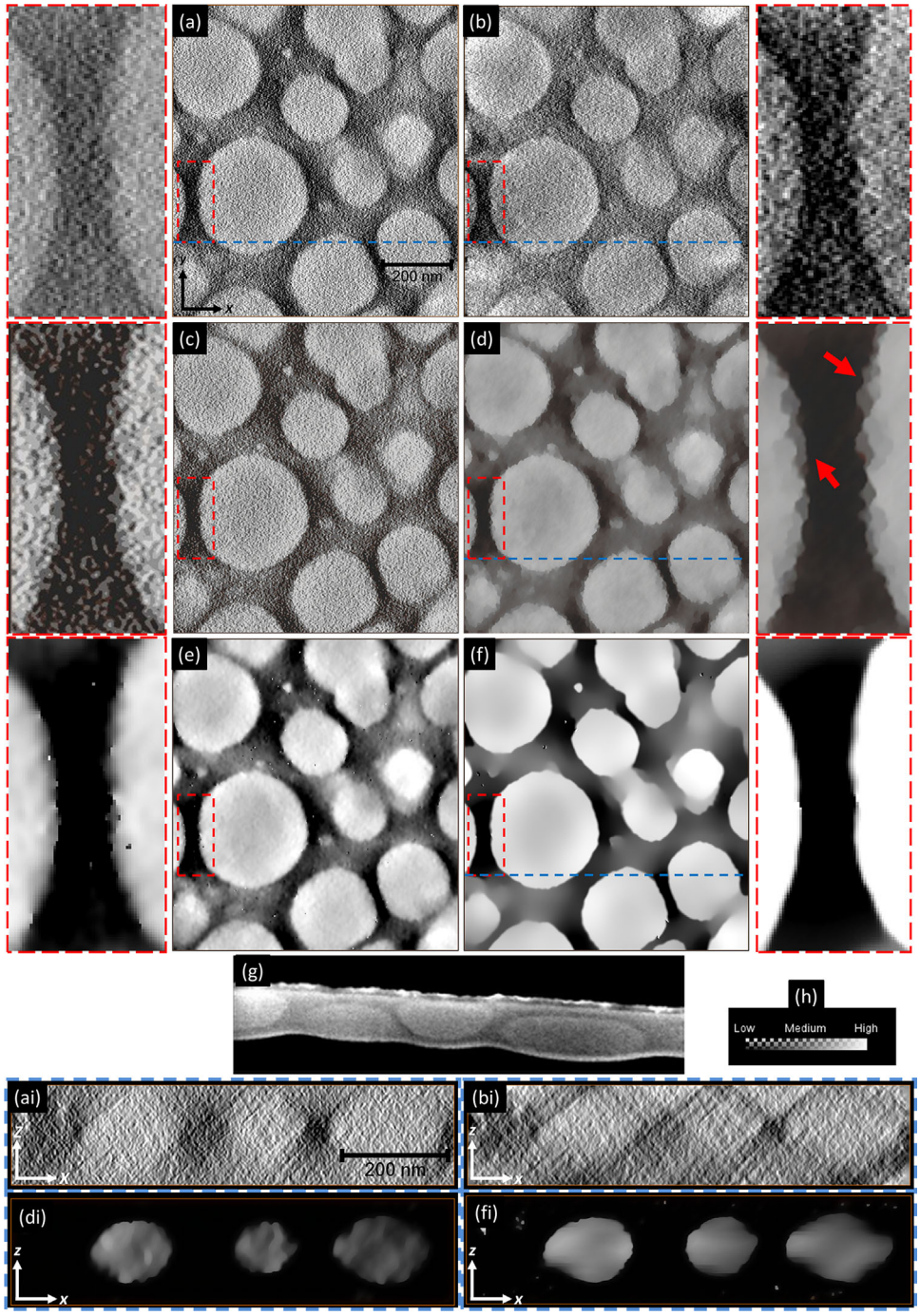

Fig 20. Reconstructions of the PTB7:$PC_{71}BM$ polymer blend solar cell sample from the experimental EFTEM tomography tilt series. (a and b) Orthoslices through WBP and SIRT reconstructions, respectively. (c and d) Orthoslices through CS-TV reconstructions with $\lambda_{l1} = 10{,}20$ respectively and $\lambda_{TV} = 10$ for both. (e and f) Orthoslices through DLET reconstructions with $\varepsilon = 2.5 \times 10^5$ and 1.3×10^4, respectively. (g) cross-sectional TEM image of the PTB7:$PC_{71}BM$ photoactive layer of the solar cell sample. See Fig. S1 in [Alekseev *et al.*, 2015] for other views. (h) Colour map for the Orthoslices. (ai and bi) XZ Orthoslices through the dotted blue line indicated on WBP and SIRT reconstructions in (a and b), respectively. (di) XZ Orthoslice through the dotted blue line indicated on the CS-TV reconstruction in (b). (fi) XZ Orthoslice through the dotted blue line indicated on the DLET reconstruction in (f).

Figure 20(di and fi), the elongation of the $PC_{71}BM$-rich domains is reduced using both CS-TV and DLET. However, we noted that different tuning for CS-TV parameters will affect the size of $PC_{71}BM$ domains, while in the case of DLET, the only effect that was noted is the blurring of the boundaries of the domains when increasing the ε above the error levels. Figure 21 shows the corresponding volume renderings of reconstructed volumes. The volume rendering display windows were limited with alpha value (transparency) reduced until the true signal from the object(s) prevailed over the background intensity.

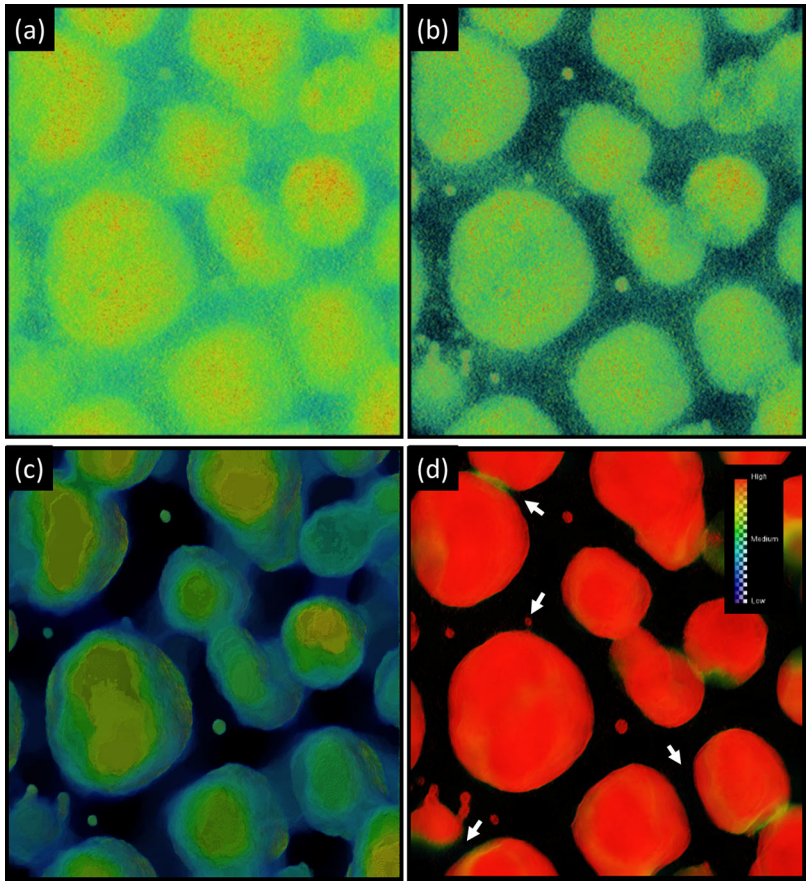

Fig 21. Direct Volume Rendering—3D perspective view voxel Z-projections of (a) WBP, (b) SIRT, (c) CS-TV and (d) DLET.

The experimental data results (Figs. 20, 21, and Table VI) are consistent with many of the features in the simulation studies. DLET was able to produce clear reconstructions with decreased streaking and missing wedge artifacts. The denoising capability was also demonstrated without sacrificing sharp details or introducing new artifacts. Also, DLET results maintained a higher SNR and contrast compared to other methods in this study, which maintains separation between different independent features (as indicated by the arrows in Fig. 21(d)). When using a value of ε that was far too large, the DLET processing will be faster; however, this will result in some loss of genuine signal as the sparse coding step of DLET will tend to approximate each patch in the reconstructed image with large errors, which can introduce some blurring. On the other hand, setting ε too small or below the noise level will increase the computing time and will add more noise and artifacts to the reconstructed image.

Conclusion and Future Work

In this article, a sparse reconstruction technique is introduced for ET that incorporates prior information through adaptive dictionaries. The dictionary is learned during the reconstruction process in a way that leads to a sparser representation of the underlying image. The proposed technique was tested using both simulated phantoms and experimental data that are known to be difficult to reconstruct using traditional (non-sparse) techniques such as the well-used TV method in compressed sensing. Reconstruction results validate its efficiency in both noiseless and noisy cases and yield an improved reconstruction quality with fast convergence. The improvement is compared with other techniques that are based on fixed sparsifying transforms in ET and with the conventional algorithms. The proposed method enables the recovery of high-fidelity information without the need to worry about which sparsifying transform to select or whether the images used strictly follow the pre-conditions of a certain transform (e.g., strictly piecewise constant for TV). This also avoids artifacts that can be introduced by specific sparsifying transforms (e.g., staircase artifacts from TV). In our future work, we aim to apply the proposed reconstruction method to other types of experimental tilt series acquired using different imaging modes in the transmission electron microscope. We also plan to investigate other optimal subsampling strategies using the signal structure in ET, since recent findings by Adcock *et al.* (2013) suggest that the optimal sampling

strategy depends not just on the overall sparsity of the signal, but also on its structure. We believe that by leveraging not just sparsity, but also structure, the CS-based ET reconstruction algorithms can be further improved so that more information can be recovered from fewer measurements. Finally, we plan to investigate the possibility of further improving the DLET algorithm for the case of extremely undersampled data via the use of separative sparse representation (SSR) which involves training an additional global dictionary to characterize artifact components. We believe that SSR can be effective as it will add a discriminative nature to the learned dictionary and exclude sparse coefficients that correspond to artifact atoms. This will enhance the dictionary's ability to characterise artifacts that are often hard to suppress without introducing blurring effects such as streak artifacts.

References

Adcock B, Hansen A, Poon C, Roman B. 2013. Breaking the coherence barrier: asymptotic incoherence and asymptotic sparsity in compressed sensing. to be published.

Aharon M, Elad M, Bruckstein A. 2006a. K -svd: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans Image Process 54:4311–4322.

Aharon M, Elad M, Bruckstein AM. 2006b. On the uniqueness of overcomplete dictionaries, and a practical way to retrieve them. Linear Algebra Appl 416:48–67.

Al-Amoudi A, Díez DC, Betts MJ, Frangakis AS. 2007. The molecular architecture of cadherins in native epidermal desmosomes. Nature 450:832–837.

AlAfeef A, Cockshott P, MacLaren I, McVitie S. 2014. Compressed sensing electron tomography using adaptive dictionaries: a simulation study. In Journal of Physics: Conference Series 522:012021. IOP Publishing.

AlAfeef A, Cockshott P, Barges B, et al. 2015a. Ebola virus reconstruction using compressed sensing cryo-electron tomography. Portland, Oregon, USA: Microscopy and Microanalysis.

AlAfeef A, Cockshott P, MacLaren I. 2015b. Linear chemically sensitive electron tomography using DualEELS and compressed sensing. Portland, Oregon, USA: Microscopy and Microanalysis.

Alekseev A, Hedley GJ, Al-Afeef A, Ageev OA, Samuel ID. 2015. Morphology and local electrical properties of ptb7: Pc 71 bm blends. J Mater Chem A 3:8706–8714.

Alpers A, Gardner RJ, Knig S, et al. 2013. Geometric reconstruction methods for electron tomography. Ultramicroscopy 128:42–54.

Arslan I, Stach EA. 2012. Electron tomography: seeing atoms in three dimensions. Nat Mater 11:911–912.

Arslan I, Yates T, Browning N, Midgley P. 2005. Embedded nanostructures revealed in three dimensions. Science 309:2195–2198.

Baraniuk, R. and Steeghs, P. 2007. Compressive radar imaging. In Radar Conference, 2007 IEEE, pages 128–133. IEEE.

Batenburg K, Bals S, Sijbers J, et al. 2009. 3d imaging of nanomaterials by discrete tomography. Ultramicroscopy 109:730–740.

Baumeister W, Grimm R, Walz J. 1999. Electron tomography of molecules and cells. Trends Cell Biol 9:81–85.

Binev P, Dahmen W, DeVore R, et al. 2012. Compressed sensing and electron microscopy. USA: Springer. p 73–126.

Bobin J, Starck J-L, Ottensamer R. 2008. Compressed sensing in astronomy. IEEE J Sel Topics Signal Process 2:718–726.

Candès EJ, Eldar YC, Needell D, Randall P. 2011. Compressed sensing with coherent and redundant dictionaries. Appl Comput Harmonic Anal 31:59–73.

Candès EJ, Romberg J, Tao T. 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans Inf Theory 52:489–509.

Candès EJ, Wakin MB. 2008. An introduction to compressive sampling. IEEE Signal Process Mag 25:21–30.

Chan TF, Esedoglu S, Park F, Yip MH. 2005. Recent developments in total variation image restoration. In Handbook of Mathematical Models in Computer Vision. Springer Verlag. p 17–33.

Chandler DM, Hemami SS. 2007. Vsnr: a wavelet-based visual signal-to-noise ratio for natural images. IEEE Trans Image Process 16:2284–2298.

Chen SS, Donoho DL, Saunders MA. 1998. Atomic decomposition by basis pursuit. SIAM J Sci Comput 20:33–61.

Cockshott WP, Tao Y, Ao G, Balch P, Briones AM, Daly C. 2003. Confocal microscopic image sequence compression using vector quantization and three-dimensional pyramid. Scanning 25:247–256.

Cotter SF, Rao BD. 2002. Sparse channel estimation via matching pursuit with application to equalization. IEEE Trans Commun 50:374–377.

Cramer H, Wold H. 1936. Some theorems on distribution functions. J London Math Soc 1:290–294.

Davenport MA, Duarte MF, Eldar YC, Kutyniok G. 2012. Introduction to compressed sensing. In: Eldar YC, Kutyniok G, editors. Compressed sensing. Cambridge University Press. Cambridge Books Online. p 1–64.

De Rosier D, Klug A. 1968. Reconstruction of three dimensional structures from electron micrographs. Nature 217:130–134.

Deans S. 1983. The radon transform and some of its applications. NY: John Wiley Sons.

Donoho D L. 2006. Compressed sensing. IEEE Trans Inf Theory 52:1289–1306.

Donoho DL, Elad M, Temlyakov VN. 2006. Stable recovery of sparse overcomplete representations in the presence of noise. IEEE Trans Inf Theory 52:6–18.

Donoho DL, Tanner J. 2005. Neighborliness of randomly projected simplices in high dimensions. Proc Natl Acad Sci USA 102:9452–9457.

Duarte MF, Davenport MA, Takhar D, et al. 2008. Single-pixel imaging via compressive sampling. IEEE Signal Process Mag 25:83–91.

Duarte MF, Eldar YC. 2011. Structured compressed sensing: From theory to applications. IEEE Trans Image Process 59:4053–4085.

Egerton RF. 2011. Electron energy-loss spectroscopy in the electron microscope. 3nd ed. US: Springer.

Elad M, Aharon M. 2006. Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans Image Process 15:3736–3745.

Elad M. 2010. The quest for a dictionary. New York: Springer.

Elad M, Figueiredo MA, Ma Y. 2010. On the role of sparse and redundant representations in image processing. Proceedings of the IEEE 98:972–982.

Engan K, Aase SO, Hakon Husoy J. 1999. Method of optimal directions for frame design. In Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on, volume 5, pages 2443–2446. IEEE.

Fornasier M, Rauhut H. 2011. Compressive sensing. In Handbook of mathematical methods in imaging. New York: Springer. p 187–228.

Foucart S, Rauhut H. 2013. A mathematical introduction to compressive sensing. Birkhäuser, Boston: Appl. Numer. Harmon. Anal.

Frank J. 2006. Electron tomography: methods for three-dimensional visualization of structures in the cell. New York: Springer.

Frank J. 2010. Electron tomography: methods for three-dimensional visualization of structures in the cell. New York: Springer.

Gan L. 2007. Block compressed sensing of natural images. In Digital Signal Processing, 2007 15th International Conference on, pages 403–406. IEEE.

Gilbert PFC. 1972. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. ii. direct methods. Proceedings of the Royal Society of London. Series B. Biol Sci 182:89–102.

Goris B, Van den Broek W, Batenburg K, Heidari Mezerji H, Bals S. 2012. Electron tomography based on a total variation minimization reconstruction technique. Ultramicroscopy 113:120–130.

Gorodnitsky IF, Rao BD. 1997. Sparse signal reconstruction from limited data using focuss: a re-weighted minimum norm algorithm. IEEE Trans Image Process 45:600–616.

Gross D, Liu Y-K, Flammia ST, Becker S, Eisert J. 2010. Quantum state tomography via compressed sensing. Phys Rev Lett 105:150401.

Herman GT. 2009. Fundamentals of computerized tomography: image reconstruction from projections. Springer.

Hofer F, Warbichler P, Grogger W. 1995. Imaging of nanometer-sized precipitates in solids by electron spectroscopic imaging. Ultramicroscopy 59:15–31.

Hofer F, Grogger W, Kothleitner G, Warbichler P. 1997. Quantitative analysis of EFTEM elemental distribution images. Ultramicroscopy 67:83–103.

Kaipio J, Erkki S. 2006. Statistical and computational inverse problems. Vol. 160. New York: Springer-Verlag.

Kak AC, Slaney M. 1988. Principles of computerized tomographic imaging. New York: IEEE press.

Kawase N, Kato M, Nishioka H, Jinnai H. 2007. Transmission electron microtomography without the missing wedge for quantitative structural analysis. Ultramicroscopy 107:8–15.

Larson PE, Hu S, Lustig M, et al. 2011. Fast dynamic 3d mr spectroscopic imaging with compressed sensing and multiband excitation pulses for hyperpolarized 13c studies. Magn Reson Med 65:610–619.

Leary R, Saghi Z, Midgley PA, Holland DJ. 2013. Compressed sensing electron tomography. Ultramicroscopy 131:70–91.

Liu B, Yu H, Verbridge SS, Sun L, Wang G. 2014. Dictionary-learning-based reconstruction method for electron tomography. Scanning 36:377–383.

Lloyd S. 1982. Least squares quantization in pcm. IEEE Trans Inf Theory 28:129–137.

Lucic V, Förster F, Baumeister W. 2005. Structural studies by electron tomography: from cells to molecules. Annu Rev Biochem 74:833–865.

Lustig M, Donoho D, Pauly JM. 2007a. Sparse mri: The application of compressed sensing for rapid mr imaging. Magnetic resonance in medicine 58:1182–1195.

Lustig M, Donoho D, Pauly JM. 2007b. Sparse mri: the application of compressed sensing for rapid mr imaging. Magn Reson Med 58:1182–1195.

Lustig M, Donoho DL, Santos JM, Pauly JM. 2008. Compressed sensing mri. IEEE Signal Process Mag 25:72–82.

Mairal J, Bach F, Ponce J, Sapiro G. 2010. Online learning for matrix factorization and sparse coding. J Mach Learn Res 11:19–60.

Mailhé B, Lesage S, Gribonval R, Bimbot F, Vandergheynst P. 2008. Shift-invariant dictionary learning for sparse representations: extending K-SVD. In the European Signal Processing Conference EUSIPCO'08, Lausanne, Switzerland. p 1–5.

Mairal J, Elad M, Sapiro G. 2008. Sparse representation for color image restoration. IEEE Trans Image Process 17:53–69.

Marziliano P, Dufaux F, Winkler S, Ebrahimi T. 2004. Perceptual blur and ringing metrics: application to JP EG2000. Image Commun Signal Process 19:163–172.

Midgley P, Weyland M. 2003. 3d electron microscopy in the physical sciences: the development of z-contrast and eftem tomography. Ultramicroscopy 96:413–431.

Midgley PA, Dunin-Borkowski RE. 2009. Electron tomography and holography in materials science. Nat Mater 8:271–280.

Monsegue N, Jin X, Echigo T, Wang G, Murayama M. 2012. Three-dimensional characterization of iron oxide (a-fe2o3) nanoparticles: application of a compressed sensing inspired reconstruction algorithm to electron tomography. Microsc Microanal 18:1362–1367.

Mostofi Y. 2011. Compressive cooperative sensing and mapping in mobile networks. IEEE Trans Mobile Comput 10:1769–1784.

Natarajan BK. 1995. Sparse approximate solutions to linear systems. SIAM J Comput 24:227–234.

Natterer F, Ẅubbeling F. 2001. Mathematical methods in image reconstruction. Philadelphia: SIAM.

Nooruddin FS, Turk G. 2003. Simplification and repair of polygonal models using volumetric techniques. IEEE Trans Vis Comput Graph 9:191–205.

Olshausen BA. 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381:607–609.

Olshausen BA, Field DJ. 1997. Sparse coding with an overcomplete basis set: a strategy employed by v1. Vision research 37:3311–3325.

Otsu N. 1975. A threshold selection method from gray-level histograms. Automatica 11:23–27.

Peyré G. 2011. A review of adaptive image representations. IEEE J Selected Topics Signal Process 5:896–911.

Protter M, Elad M. 2009. Image sequence denoising via sparse and redundant representations. IEEE Trans Image Process 18:27–35.

Radermacher M. 1988. Three-dimensional reconstruction of single particles from random and nonrandom tilt series. J Electron Microsc Tech 9:359–394.

Radon J. 1936. Ü ber die bestimmung von funktionen durch ihre integralwerte 294 längs gewisser mannigfaltigkeiten, ber. Verh. Säch. Akad. Wiss. Leipzig 295:262–277.

Ramachandran G, Lakshminarayanan A. 1971. Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of fourier transforms. PNAS 68:2236–2240.

Rauhut H, Schnass K, Vandergheynst P. 2008. Compressed sensing and redundant dictionaries. IEEE Trans Inf Theory 54:2210–2219.

Rice, U. Compressive sensing resources @ONLINE. April 2014. Url: http://dsp.rice.edu/cs.

Robinson CV, Sali A, Baumeister W. 2007. The molecular sociology of the cell. Nature 450:973–982.

Saghi Z, Holland DJ, Leary R, et al. 2011. Threedimensional morphology of iron oxide nanoparticles with reactive concave surfaces. a compressed sensing-electron tomography (cs-et) approach. Nano Lett 11:4666–4673.

Sayood K. 2002. Statistical evaluation of image quality measures. J Electron Imaging 11:206–223.

Sallee P, Olshausen B. 2002. Learning sparse multiscale image representations. In Advances Neural Information Processing Systems. Vancouver, BC, Canada: MIT Press. 15:1327–1334.

Schaffer B, Grogger W, Kothleitner G. 2004. Automated spatial drift correction for eftem image series. Ultramicroscopy 102:27–36.

Scott M, Chen C-C, Mecklenburg M, et al. 2012. Electron tomography at 2.4-angstrom resolution. Nature 483:444–447.

Smith, D. and Welch, E. (2011). Non-sparse phantom for compressed sensing mri reconstruction. In International Society for Magnetic Resonance in Medicine 19th Scientific Meeting-ISMRM, 11, 2845.

Starck J-L, Elad M, Donoho D. 2004. Redundant multiscale transforms and their application for morphological component separation. Adv Imaging Electron Phys 132:287–348.

Stevens A, Yang H, Carin L, Arslan I, Browning ND. 2013. The potential for bayesian compressive sensing to significantly reduce electron dose in high-resolution stem images. J Electron Microsc 63:41–51.

Tao Y, Cockshott W. 2004. Three-dimensional microscopic image coding by finite-state vector quantization in an enhanced image pyramid. Proc SPIE 5370:1895–1905.

Tibshirani R. 1996. Regression shrinkage and selection via the lasso. J R Statistical Soc Series B Stat Methodol 58:267–288.

Tosic I, Frossard P. 2011. Dictionary learning. IEEE Signal Process Mag 28:27–38.

Tropp JA, Gilbert AC. 2007. Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inf Theory 53:4655–4666.

Tropp JA, Wright SJ. 2010. Computational methods for sparse solution of linear inverse problems. Proc IEEE 98:948–958.

Tsaig Y, Donoho DL. 2006. Extensions of compressed sensing. Signal Process 86:549–571.

van Bavel SS, Loos J. 2010. Volume organization of polymer and hybrid solar cells as revealed by electron tomography. Adv Funct Mater 20:3217–3234.

van Den Berg E, Friedlander MP. 2008. Probing the pareto frontier for basis pursuit solutions. SIAM J Sci Comput 31:890–912.

van den Broek W, Rosenauer A, Goris B, et al. 2012. Correction of non-linear thickness effects in haadf stem electron tomography. Ultramicroscopy 116:8–12.

Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13:600–612.

Weyland M, Midgley PA. 2003. Extending energy-filtered transmission electron microscopy (eftem) into three dimensions using electron tomography. Microsc Microanal 9:542–555.

Williams DB, Carter CB. 2009. Transmission electron microscopy: a textbook for materials science. NY: Spinger.

Wipf DP, Rao BD. 2004. Sparse bayesian learning for basis selection. IEEE Trans Image Process 52:2153–2164.

Wright J, Ma Y, Mairal J, et al. 2010. Sparse representation for computer vision and pattern recognition. Proceedings of the IEEE 98:1031–1044.